

WO 99/32147

PCT/US98/27364

1

pVR 1012-GP(IC)

Sequence Listing ID No: 1

General Description

DNA pVR 1012-GP(IC)
Local object
Created: 09/14/98 04:17PM
Last Modification Date: ? (no data)
length: 7003 bp
storage type: Basic
form: Circular

Comments

Restriction Map

BglII: 1 site AGATCT
TCTAGA

Clal: 1 site ATCCAT
TAGCTA

DrallI: 1 site CACNNNGTG
GTGNNNCAC

EcoRV: 1 site GATATC
CTATAG

HindIII: 1 site AAGCTT
TTCGAA

HpaI: 1 site GTTAAC
CAATTG

KasI: 1 site GGCGCC
CCGCGG

KpnI: 1 site GGTACC
CCATGG

NarI: 1 site GGCGCC
CCGCGG

PmlI: 1 site CACGTG
GTGCAC

PstI: 1 site CTCGAG
GAGCTC

PvuI: 1 site CGATCG
GCTAGC

SacII: 1 site CCGCGG
GGCGCC

Sall: 1 site GTCGAC
CAGCTG

XmnI: 1 site GAANNNTTC
CTTNNNAAG

EcoRI: 2 sites GAATTC
CTTAAG

NcoI: 2 sites CCATGG
GGTACC

NdeI: 2 sites CATATG
GTATAC

SphI: 2 sites GCATGC
CGTACG

XhoI: 2 sites CTCGAG
GAGCTC

BamHI: 3 sites CGATCC
CCTAGG

WO 99/32147

2

PCT/US98/27364

BclI: 3 sites TCATCA
 ACTAGT

Functional Map**CDS (4 signals)****CMV IE 5' UT**

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4020 End: 4572

Kan r

Start: 6068 End: 6690 (Complementary)

Misc_feature (2 signals)**CMV enhancer**

Start: 248 End: 885

GP(IC)

Start: 1870 End: 4019

Annotations

```

1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
   AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC
.....
51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
   CTCTGCCAGT GTCGAACAGA CATTGCCCTA CGGCCCTCGT CTGTTGCGGC
.....
101 TCAGCGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACCGGCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTCTA CTGAGAGTCC ACCATATGCC GTGTGAAATA
   GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT
.....
201 CCGCACAGAT CCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
   GCGGTGTCTA CGCATTCTTC TTTTATGGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTATA TTGGCTCATG
   AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGGTTGTAAT GCGCGTACAA CTGTAACATA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCAAGGCC CATATATGCA GTTCCGCGTT
   TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA
.....
401 ACATAACTTA CCGTAAATGG CCCGCCCTGGC TGACCGCCCA ACGACCCCGG
   TGTATTGAAT GCCATTACCG GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
   GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTATCCCT
.....
501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTG ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCACTAG TTACATAGT ATACGGTTCA TCGGGGGGAT AACTGCACTT
.....
601 TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATTT ACCGGCGGGA CCGTAATACG GGTCAATGAT TGGAAATACCC
.....

```

NcoI

```

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TGAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
   CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG
.....
751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCTATAA GGTTCAGAGG TGGGGTAAC TCAAGTTACCC TCAAACAAA
.....
801 GGCACCAAAA TCAACGGGAC TTTCAAAAT GTCGTAACAA CTCGCCCCCA
   CCGTGCTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
.....

```

851 TTGACGCAAA TCGGCGGTAG GCGGTACCG TGGGAGGTCT ATATAAGCAG
AACITGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCGGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GTCCGGAGGC CCCGGCCCTT

1001 CCGTGCAATTG GAACCGGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCAATGCGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGAATAAC TGGTGAGGGG

1201 TATTGGTGAC CATACTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTGATAGAG ATAACCGATA TACGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
TGCCTGAGAC ATAAAAATGT CCTACCCCAAG GGTAAATAAT AAATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCTGCC CGCAGTTTAT ATAAACATA
GTCTATATGT TGTTCGGCA GGGGACACGG GCGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACGCCAA TCTCGGTAC GTGTCCGGA CATGGGCTCT
CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGCCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
AGAGGCCATC GCCGCCCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCGCCGAGTA CCAGCGAGCC GTCCAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTCCCG CACAAGCCCG
TGAATCCGTG TCGTGTACG GGTGTTGGTG GTACACGGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCAGG
ACCGCCATCC CATAACAGCA CTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GGTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTTAACGGT GGAGCCCACT GTAGTCTGAG CAGTACTCGT TGCTGCCCGG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACCACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGCTGACA AGGAAAGGTA

SaiI

NcoI PstI PmlI BclI EcoRV

1851 GGGTCTTTTC TGCAGTCACC GTCGTGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

EcoRI

1901 CCGCGCGGCC GCTCTAGAAT TCTCTAATCA CAGTCATCAT GCGAGCGTCA
GCGCGCGCCG CGAGATCTTA AGAGATTAGT GTCAGTAGTA CCCTCGCAGT

1951 GGGATTCTGC AATTGCCCGG TGAGCGCTTC AGGAAAACAT CTTTCTTTGT
CCCTAAGACG TTAACGGGGC ACTCGCGAAG TCCTTTTGTA GAAAGAAACA

2001 TTGGGTAATA ATCCTATTCC ATAAAGTCTT TTCAATCCCG TTGGGGGTTG
AACCCATTAT TAGGATAAGG TATTTACAAA AAGTTAGGGC AACCCCAAC

2051 TACACAACAA TACCCTACAA GTGAGTGATA TTGACAAGTT TGTGTGCCGA
ATGTGTTCTT ATGGGATGTT CACTCACTAT AACTGTTCAA ACACACGGCT

2101 CACAAACTCT CTTCAACTAG CCAATTGAAG TCAGTCGGGT TGAACCTGGA
CTGTTTGAGA GAAGTTGATC GGTAACTTC AGTCAGCCCA ACTTGAACCT

2151 GGGCAATGGA GTAGCAACTG ATGTACCAAC GGCAACCAA AGATCGGGTT
CCCCTTACCT CATCGTTGAC TACATGGTTG CCGTTGGTTT TCTACCCCAA

2201 TTCGAGCTGG TGTCCACCA AAGGTGGTAA ATTACGAAGC TGGAGAATGG
AAGCTCGACC ACAAGGTGGT TTCCACCATT TAATGCTTCG ACCTCTTACC

2251 GCTGAGAACT GTTATAACCT GGCTATAAAG AAAGTTGATG GTAGTGAGTG
CGACTCTTGA CAATATTGGA CCGATATTTC TTTCAACTAC CATCACTCAC

2301 CCTACCAGAA GCCCCTGAGG GAGTGAGGGA TTTTCCCCGT TGCCGCTATG
GGATGCTCTT CGGGGACTCC CTCCTCCCT AAAAGGGGCA ACGGCGATAC

2351 TACACAAAGT CTCAGCAACT GGACCATGCC CAGGAGGACT CGCCTTTCAC
ATGTGTTTCA GAGTCCTTGA CCTGGTACGG GTCCTCCTGA GCGGAAAGTG

2401 AAAGAAGGAG CCTTCTTCCT GTATGACCGA CTCGCATCAA CAATCATTTA
TTTCTTCCTC GGAAGAAGGA CATACTGGCT GAGCGTAGTT GTTAGTAAAT

2451 TCGGGGTACA ACCTTTGCCG AAGGAGTTAT TGCATTTCTG ATCTTGCCTA
AGCCCCATGT TCGAAACGGC TTCCTCAATA ACGTAAAGAC TAGAACGGAT

2501 AGGCGCGAAA GGATTTTTTC CAGTCTCCTC CATTGCATGA GCCTGCCAAC
TCCGCGCTTT CCTAAAAAG GTCAGAGGAG GTAACGTACT CGGACGGTTG

BamHI

2551 ATGACCACGG ATCCCTCCAG TTACTATCAC ACGACAACAA TAAACTACGT
TACTCGTGCC TAGGGAGGTC AATGATAGTG TGCTGTTGTT ATTTGATGCA
.....
2601 GGTGATAAAT TTTGGAACCA ACACCACAGA GTTCTGTTC CAAGTCGATC
CCAACTATTA AAACCTTGGT TGTGGTGTCT CAAAGACAAG GTTCAGCTAG
.....

XhoI

2651 ATTTGACGTA TGTGCAGCTC GAGGCAAGAT TCACACCACA ATTCCTTGTC
TAAACTGCAAT ACACGTCGAG CTCCGTTCTA AGTCTGGTGT TAAGGAACAG
.....
2701 CTCCTAAATG AAACCATCTA CTCTGATAAC CGCAGAAGTA ACACAACAGG
GAGGATTTAC TTTGGTAGAT GAGACTATTG GCGTCTTCAT TGTGTTGTCC
.....
2751 AAAACTAATC TGGAAATAA ATCCCACTGT TGATACCAGC ATGGGTGAGT
TTTGTATTAG ACCTTTTATT TAGGGTGACA ACTATGGTCC TACCCACTCA
.....
2801 GGCCTTTCTG GGAATAATA AAAACTTCAC AAAAACCTT TCAAGTGAAG
CCCGAAAGAC CCTTTTATT TTTTGAAGTG TTTTGGGAA AGTTCAC TTC
.....
2851 AGTTGTCTTT CGTACCTGTA CCAGAAACCC AGAACCAGGT CCTTGACACG
TCAACAGAAA GCATGGACAT GGTCTTTGGG TCTTGGTCCA GGAACGTGTC
.....
2901 ACAGCGACGG TCTCTCCTCC CATCTCCGCC CACAACCACG CAGGCGAAGA
TGTCGCTGCC AGAGAGGAGG GTAGAGGCGG GTCTTGGTGC GTCCGCTTCT
.....
2951 CCACAAAGAA TTGGTTTCAG AGGATTCCAC TCCAGTGGTT CAGATGCAAA
CGTGTTTCTT AACCAAAGTC TCCTAAGGTG AGGTCACCAA GTCTACGTTT
.....
3001 ACATCAAGSG AAAGGACACA ATGCCAACCA CAGTGACGGG TGTACCAACA
TGTACTTCCC TTTCTGTGT TACGGTTGGT GTCAGTCCCC ACATGTTTGT
.....

BclI

3051 ACCACACCCT CTCCATTTCC AATCAATGCT CGCAACACTG ATCATACCAA
TGGTCTGGGA GAGGTAAAGG TTAGTTACCA GCGTTGTGAC TAGTATGGTT
.....
3101 ATCATTATC GGCCTGGAGG GGCCCAAGA AGACCACAGC ACCACACAGC
TAGTAAATAG CCGGACCTCC CCGGGTTCT TCTGGTGTG TGGTGTGTCG
.....
3151 CTGCCAAGAC CACCAGCCAA CCAACCAACA GCACAGAATC GACGACACTA
GACGGTTCTG GTGGTGGGT GGTGGTTGT CGTGTCTTAG CTGCTGTGAT
.....
3201 AACCCAACAT CAGAGCCCTC CAGTAGAGGC ACGGACCAT CCAGCCCCAC
TTGGGTTGTA CTCTCGGGAG GTCATCTCCG TCCCTGGTA GGTGGGGTG
.....
3251 GGTCCCAAC ACCACAGAAA GCCACGCCGA ACTTGGCAAG ACAACCCAA
CCAGGGGTTG TGGTGTCTTT CCGTGCGGCT TGAACCGTTC TGTGGGGTT
.....
3301 CCACACTCCC AGAACAGCAC ACTGCCGCCA GTGCCATTCC AAGAGCCGTG
GGTGTGAGGG TCTTCTCTGT TGACGGCGGT CACGGTAAGG TTCTCGGCAC
.....
3351 CACCCCGACG AACTCAGTGG ACCTGGCTTC CTGACGAACA CAATACGGGG
GTGGGGCTCC TTGAGTCACC TGGACCGAAG GACTGCTTGT GTTATGCCCC
.....

BamHI

3401 GGTGACAAAT CTCCTGACAG GATCCAGAAG AAAGCGAAGG GATGTCACTC
 CCACCTGTTTA GAGGACTGTC CTAGGTCTTC TTTCGCTTCC CTACAGTGAG

 3451 CCAATACACA ACCCAAATGC AACCCAAACC TGCCTATTG CACAGCCTTG
 GGTATATGTGT TGGGTTTACG TTGGGTTTGG ACGTGATAAC CTGTGCGAAC

 3501 GATGAGGGTG CTCCCATAGG TTTAGCCTGG ATACCATACT TCGGGCCAGC
 CTACTCCAC GACGGTATCC AAATCGGACC TATGGTATGA AGCCCGGTCG

 3551 AGCTGAGGGA ATTTACACTG AAGGCATAAT CGAGAATCAA AATGGATTGA
 TCGACTCCCT TAAATGTGAC TTCCGTATTA CCTCTTAGTT TTACCTAACT

 3601 TCTGTGGATT GAGGCAGCTG GCCAACGAAA CGACACAAGC TCTTCAATG
 AGACACCTAA CTCCGTCGAC CGGTTGCTTT GCTGTGTTCC AGAAGTTAAC

 3651 TTCTTAAGGG CAACTACTGA GTTCCGTACA TTCTCTATAC TAAATCGGAA
 AAGAATTCCC GTTGATGACT CAACGCATGT AAGAGATATG ATTTAGCCTT

 3701 AGCAATACAC TTCTTGCTCC AAAGATGGGG AGGAACATGT CACATCTAG
 TCGTTATCTG AAGAACGAGG TTTCTACCCC TCCTTGATCA GTGTANGATC

 3751 GGCCTGATTG TTGCATTGAA CCCCAAGATT GGACCAAAA TATCACTGAT
 CCGGACTAAC AACGTAACCTT GGGGTTCTAA CCTGGTTTTT ATAGTGACTA

BclI

3801 AAAATTGATC AAATAATCCA TGACTTTGTC GATAATAATC TTCCAAATCA
 TTTAACTAG TTTATTAGGT ACTGAAACAG CTATTATTAG AAGGTTTAGT

 3851 GAATGATGGC AGCAACTGGT GGAATGGATG GAAACAATGG GTTCCTGCTG
 CTTACTACCG TCGTTGACCA CCTGACCTAC CTTTGTTACC CAAGGACGAC

 3901 GAATAGGAAT CACAGGAGTA ATCATTCCTA TTATTGCTTT GCTGTGCATT
 CTTATCCTTA GTGTCTCAT TAGTAACGAT AATAACGAAA CGACACGTAA

EcoRI

3951 TGCAAATTCA TGCTTTGAAC TAATATAGCA TCATACTTTA GAATTCCTAGA
 ACGTTTAAGT ACGAACTTG ATTATATCGT AGTATGAAAT CTTAAGATCT

NarIKasIBamHI BglII

4001 CCAGGCGCCT GGATCCAGAT CTGCTGTGCC TTCTAGTTCC CAGCCATCTG
 GGTCCGCGGA CCTAGGTCTA GACGACACGG AAGATCAACG GTCGGTAGAC

 4051 TTGTTTGCCC CTCCCCGTG CTTTCTTGA CCTGGAAGG TGCCACTCCC
 AACAAACGGG GAGGGGGCAC GGAAGGAAT GGGACCTTC ACCGTGAGGG

 4101 ACTGTCCTTT CTAATAAAAA TGAGGAAAT GCATCGCATT CTCTGAGTAG
 TGACAGGAAA GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC

 4151 GTGTCAATCT ATTCTGGGGG GTGGGGTGGG GCAGCACAGC AAGGGGGAGG
 CACAGTAAGA TAAGACCCCC CACCCACCCC CGTCGTGTCG TTCCCCCTCC

	<u>SphI</u>		<u>KpnI</u>	
4201	ATTGGAAGA	CAATAGCAGG	CATGCTGGGG	ATGCGGTGGG
	TAACCCCTTCT	GTTATCGTCC	GTACGACCCC	TACGCCACCC
			GAGATACCCA	
.....				
	<u>KpnI</u>			
4251	ACCCAGCTGC	TGAAGAATTG	ACCCGGTTTC	TCCTGGGCCA
	TGGGTCCACG	ACTTCTTAAC	TGGGCCAAGG	AGGACCCCGT
			CTTCTTCGT	
.....				
4301	CCCACATCCC	CTTCTCTGTG	ACACACCCCTG	TCCACGCCCC
	CCGTGTAGGG	GAAGAGACAC	TGTGTGGGAC	AGGTGCGGGG
			ACCAAGAATC	
.....				
4351	TCCAGCCCC	ACTCATAGGA	CACTCATAGC	TCAGGAGCGC
	AAGGTGCGGG	TGAGTATCCT	GTGAGIATCG	AGTCTCTCCG
			AGGCGGAAGT	
.....				
4401	ATCCACCCC	CTAAAGTACT	TGGAGCGGTC	TCTCCCTCCC
	TAGGGTGGGC	GATTTTCATGA	ACCTCGCCAG	AGAGGGAGGG
			AGTAGTCGGG	
.....				
4451	ACCAAACCAA	ACCTAGCCTC	CAAGAGTGGG	AAGAAATTAA
	TGGTTTGGTT	TGGATCGGAG	GTTCTCACCC	TTCTTTAATT
			TCGTTCTATC	
.....				
4501	GCTATTAAAT	GCAGAGCGAG	AGAAAATGCC	TCCAACATGT
	CGATAATTCA	CGTCTCCCTC	TCTTTACGG	AGGTTGTACA
			CTCCTTCATT	
.....				
	<u>XbaI</u>			
4551	TGAGAGAAAT	CATAGAAATT	CTTCCGCTTC	CTCGCTCACT
	ACTCTCTTTA	GTAATCTTAA	GAAGGCCAAG	GAGCGAGTGA
			CTGAGCGACG	
.....				
4601	CCTCGGTCGT	TGGGTGCGG	CGAGCGGTAT	CAGCTCACTC
	CGAGCCAGCA	AGCCGACGCC	CCTCGCCATA	GTGAGTGAG
			TTTCCCCAT	
.....				
4651	ATACGTTAT	CCACAGAATC	AGGGGATAAC	GCAGGAAAGA
	TATGCCAATA	GGTGTCTTAG	TCCCTATTG	CGTCTTTCT
			TGTACACTCG	
.....				
4701	AAAAGGCCAG	CAAAAGGCCA	GGAACCGTAA	AAAGGCCGCG
	TTTTCCGGTC	GTTTCCGGT	CCTTGGCATT	TTTCCGGCGC
			AACGACCGCA	
.....				
4751	TTTTCCATAG	GCTCCGCCCC	CCTGACGAGC	ATCACAAAAA
	AAAAGGTATC	CGAGGCGGGG	GGACTGCTCG	TAGTGTTTT
			AGCTCCGAGT	
.....				
4801	AGTCAGAGGT	GCGGAAACCC	GACAGGACTA	TAAAGATACC
	TCAGTCTCCA	CCGCTTTGGG	CTGTCTGAT	ATTTCTATGG
			TCCGCAAAGG	
.....				
4851	CCCTGGAAGC	TCCCTCGTGC	GCTCTCTGT	TCCGACCCCTG
	GGGACCTTCG	AGGGAGCAGG	CGAGAGGACA	AGGCTGGGAC
			GGCGAATGGC	
.....				
4901	GATACCTGTC	CGCCTTTCTC	CCTTCGGGAA	GCCTGGCGCT
	CTATGCACAG	GCGGAAAGAG	GGAAGCCCTT	CGCACCCCGA
			AAGAGTTACC	
.....				
4951	TCACGCTGTA	GGTATCTCAG	TTCGGTGTAG	GTGTTGCTG
	AGTCCGACAT	CCATAGAGTC	AAGCCACATC	CAGCAAGCGA
			GGTTCGACCC	
.....				
5001	CTGTCTCCAC	GAACCCCCCG	TTCAGCCCGA	CCGCTGCGCC
	GACACACGTG	CTTGGGGGGC	AAGTCGGGCT	GGCGACGCGG
			AATAGGCCAT	
.....				


```

5051 ACTATCGTCT TGAGTCCAAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA
    TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT
.....
5101 GCAGCCACTG GTAACAGGAT TAGCAGACCG AGGTATGTAG GCGGTGCTAC
    CGTCGGTGAC CATGTGCTTA ATCGTCTCGC TCCATACATC CGCCACCATG
.....
5151 AGAGTTCTTG AAGTGGTGGC CTAACCTACG CTACACTAGA AGGACAGTAT
    TCTCAAGAAC TTCACCACCG GATTGATGCC CATGTGATCT TCCTGTCATA
.....
5201 TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGACTTGGT
    AACCATAGAC GCGACACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA
.....
5251 AGCTCTTGAT CCGGCAAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
    TCGAGAACTA GCGCGTTTGT TTGGTGGCGA CCATCGCCAC CAAAAAACA
.....
5301 TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
    AACGTTGCTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA
.....
5351 TGACTTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA
    ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT
.....
5401 GGGATTTTGG TCATGAGATT ATCAAAAAGC ATCTTCACCT AGATCCTTTT
    CCTTAAACC ACTACTCTAA TAGTTTTTCC TAGAAGTCCA TCTAGGAAA
.....
5451 AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAACTT
    TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA
.....
5501 GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC
    CCAGACTGTC AATGGTTACG AATTAGTCAC TCCGTGSATA GAGTCGCTAG
.....
5551 TGCTATTTTC GTTCATCCAT AGTTGCCTGA CTCCGGGGGG GGGGGCGGCT
    ACAGATAAAG CAAGTAGSTA TCAACGGACT GAGGCCCCCC CCCCCGCGA
.....
5601 GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCTGAATC
    CTCCAGACGG ACCACTTCTT CCACAACGAC TGAGTATGGT CCGGACTTAG
.....
5651 GCCCCATCAT CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT
    CGGGGTAGTA GGTGGGTCTT TCACTCCCTC GGTGCCAAT ACTCTCGAAA
.....
5701 GTTGTAGGTG GACCAGTTGG TGATTTTGAA CTTTGTCTTT GCCACGGAAC
    CAACATCCAC CTGGTCAACC ACTAAAATT GAAAACGAAA CGGTGCCTTG
.....
5751 GGTCTGCGTT GTCGGGAAGA TCGGTGATCT GATCCTTCAA CTCAGCAAAA
    CCAGACGCAA CAGCCCTTCT ACGCACTAGA CTAGGAAGTT GAGTCGTTTT
.....
5801 GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG CGTAATGCTC
    CAACCTAAAT AACTTGTTTC GGCGGCAGGG CAGTTCAGTC GCATTACGAG
.....
5851 TGCCAGTGTG ACAACCAATT AACCAATTCT CATTAGAAAA ACTCATCCAG
    ACGGTCACAA TGTTGGTTAA TTGGTTAAGA CTAATCTTTT TGAGTAGCTC
.....
5901 CATCAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT
    GTAGTTTACT TTGACGTTAA ATAAGTATAG TCCTAATAGT TATCGTATAA
.....
5951 TTTGAAAAAG CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC
    AAACTTTTTC GGCAAGACA TACTTCCTC TTTTGAGTGG CTCCGTCAAG
.....

```

6001 CATAGGATGG CAAGATCCTG GTATCGGTCT GCGATTCCGA CTCGTCCAAC
GTATCCTACC GTTCTAGGAC CATAGCCAGA CCTAAGGCT GAGCAGGTTG

6051 ATCAATACAA CCTATTAAAT TCCCCTCGTC AAAAATAAGG TTATCAAGTC
TAGTTATCTT GGATAATTAA AGGGGAGCAG TTTTATTCC AATAGTTCAC

HindIII

6101 AGAAATCACC ATGAGTGACG ACTCAATCCG GTGAGAATCG CAAAAGCTTA
TCCTTAGTGG TACTCACTGC TGACTTAGGC CACTCTTACC GTTTTCGAAT

6151 TGCATTTCTT TCCAGACTTG TTCAACAGGC CAGCCATTAC GCTCGTCATC
ACGTAAGAA AGGTCTGAAC AAGTTGTCCG GTCCGTAATG CGAGCAGTAG

6201 AAAATCACTC GCATCAACCA AACCGTTATT CATTCGTGAT TGCGCCTGAG
TTTTAGTGAG CGTAGTTGGT TTGGCAATAA GTAAGCACTA ACCCGGACTC

PvuI

6251 CGAGACGAAA TACGCGATCG CTGTTAAAAG GACAATTACA AACAGGAATC
GCTCTGCTTT ATGCGCTAGC GACAACTTC CTGTTAATGT TTGTCCTTAG

6301 GAATGCAACC GCGCGAGGAA CACTGCCAGC GCATCAACAA TATTTTCACC
CTTACGTTGG CCGCGTCCCT GTGACGCTCG CGTAGTTGTT ATAAAAGTGG

6351 TGAATCAGGA TATTCTTCTA ATACCTGCAA TGCTGTTTC CCGGGGATCG
ACTAGTCCT ATAAGAAGAT TATGGACCTT ACGACAAAAG GGCCCCTAGC

6402 CAGTCGTGAG TACCATGCA TCATCAGGAG TACGGATAAA ATGCTTGATG
GTCACCACTC ATTGGTACGT AGTAGTCCTC ATGCCTATTT TACGAACCTAC

6451 GTCGGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCATCTCATC
CAGCCTCTC CGTATTTAAG GCAGTCGGTC AAATCAGACT GGTAGAGTAG

6501 TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG
ACATTGTAGT AACCGTTGCG ATGGAACGG TACAAAGTCT TTGTTGAGAC

ClaI

6551 GCGCATCGGG CTTCCCATAC AATCGATAGA TTGTCCGACC TGATTGCCCG
CCCGTAGCCC GAAGGGTATG TTAGCTATCT AACAGCGTGG ACTAACGGGC

6601 ACATTATCGC GAGCCCATTT ATACCCATAT AAATCAGCAT CCATGTTGGA
TGTAATAGCG CTCGGGTAAA TATGCCATA TTTAGTCGTA GGTACAACCT

XhoI

6651 ATTTAATCGC GGCCTCGAGC AAGACGTTTC CCGTTGAATA TGGCTCATAA
TABATTAGCG CCGGAGCTCG TTCTGCAAAG GGCAACTTAT ACCGAGTATT

6701 CACCCCTTGT ATTACTGTTT ATGTAAGCAG ACAGTTTTAT TGTTTCATGAT
GTGGGGAACA TAATGACAAA TACATTGCTC TGTCAAAATA ACAAGTACTA

DraIII

6751 GATATATTTT TATCTTCTGC AATGTAACAT CAGAGATTTT GAGACACAAC
CTATATAAAA ATAGAACACG TTACATTGTA GTCTCTAAAA CTCTCTGTTG

DraIII

6801 GTEGCTTTCC CCCCCCCCCC ATTATTGAAG CATTATCAG GGTATTGTC
CACCGAAAGC GGGGGGGGGG TAATAACTTC GTAAATAGTC CCAATAACAG
.....
6851 TCATGAGCCG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG
AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTATT TGTATTATCCC
.....
6901 GTTCCGCCCA CATTTCCTCCG AAAAGTCCCA CCTGACGTCT AAGAAACCAT
CAAGGCGCGT GTAAAGGGGC TTTTCACGGT GGACTGCACA TTCTTTGGTA
.....
6951 TATTATCATG ACATTACCT ATAAAAATAG GCGTATCAG AGCCCTTTC
ATAATAGTAC TGTAAATTGA TATTTTATC CCCATAGTGC TCCGGGAAG
.....
7001 GTC
CAG
.....

pVR 1012-GP(S)

Sequence Listing ID No: 2

General Description

DNA pVR 1012-GP(S)
 Local object
 Created: 09/14/98 03:58PM
 Last Modification Date: ? (no data)
 length: 7073 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

BalI: 1 site TGGCCA
ACCGGT

BclI: 1 site TGATCA
ACTAGT

Clal: 1 site ATCGAT
TAGCTA

DraIII: 1 site CACNNGTG
GTGNNCAC

HindIII: 1 site AAGCTT
TTCGAA

KasI: 1 site GCGGCC
CCGCGG

KpnI: 1 site GGTACC
CCATGG

NarI: 1 site GCGGCC
CCGCGG

PmlI: 1 site CACGTG
GTGCAC

PvuI: 1 site CGATCG
GCTAGC

SacII: 1 site CCGCGG
GGCGCC

Sall: 1 site GTCGAC
CAGCTG

XbaI: 1 site TCTAGA
AGATCE

XmnI: 1 site GAANNNTTC
CTTNNNAAG

NdeI: 2 sites CATATG
GTATAC

EcoRV: 3 sites GATATC
CTAATG

SphI: 3 sites GCATGC
CGTACG

NcoI: 4 sites CCATGG
GGTACC

BamHI: 6 sites GGATCC
CCTAGG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4090 End: 4642

Kan r

Start: 6138 End: 6760 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(S)

Start: 1870 End: 4089

Annotations

1 TCGCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCC
 A3CGCGCAAA GCCACTACTG CCACCTTTGG AGACTGTGTA CGTCGAGGGC

 51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCC
 CTCGCGCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC

 101 TCAGGGCCCG TCAGCGGGTG TTGCGGGTG TCGGGGCTGG CTTAACTATG
 AGTCCCGCGC AGTCGCCCCAC AACCGCCCCAC AGCCCCGACC GAATTGATAC

NdeI

151 CGGCATCAGA GCAGATTGTA CTGAGACTGC ACCATATGCG GTGTGAAATA
 CCCGTAGTCT CGTCTAACAT GACTCTCAGC TGTATACGC CACACTTTAT

BalI

201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
 GCGGTGTCTA CGCATTCTTC TTTTATGGCG TAGTCTAACG GATAACCGGT

 251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
 AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
 AGGTTGTAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAATTATCA

 351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
 TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401 ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
 TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
 GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTCG GGTATCCCT

 501 CTTTCCATTG ACGTCAATGG GTGCAATAT TACGGTAAAC TGCCCACTTG
 GAAAGGTAAC TCCAGTTACC CACCTCATAA ATGCCATTG ACGGCTGAAC

NdeI

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
 CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601 TCACGGTAAA TGGCCCCGCT GGCATTATGC CCAGTACATG ACCTTATGGG
 ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGAATACCC

NcoI

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
 TGAAAGCATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

NcoI

701 GTGATCGGGT TTTGGCAGTA CATCAATGGG CGTGGATACC GGTTCAGTCT
 CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTGTT
 TCGCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAAACAAA

801 GCCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT

851 TTGACGCAAA TCGGCGGTAG GCGTGACGG TGGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGGCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTCGGGTA GGTCCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT

1001 CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGCCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTOCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTC CTATGAAAGG TAATGATTAG GTATTCTACC GAGAAACGGT

1251 CAACATATCT TATTGGCTAT ATGCCAATAC TCTGTCTTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGCTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
TGCCCTGAGAC ATAAAAATG CCTACCCAG GGTAAATAAT AATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTC ATTAAACATA
GTGTATATGT TGTTCGGGCA GGGGGCACGG CGGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
CGCACCCCTAG AGGTGCGCTT AGAGCCCATC CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCCTC
AGAGCCCATC GCCGCCCTCA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCAG
ACCGCCATCC CATACACAGA CTTTACTCG CACCTCTAAC CCGACGTC

1651 GCTGACGAG ATGGAAGACT TAAGGCAGCG GCAGAAGAG ATCCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AACAGTCACA GGTAACCTCC GTTGCCGTGC
GACTCAACAA CATAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG

1751 TGTAAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NeoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

Sall

NeoI

PmlI BclI EcoRV

1851 CCGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTCT GCACACTAGT CTATAGCGCC

SplI

EcoRV

1901 CCGCTCTAGC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT
GGCGAGATCG ATCTACGTAC GAGCTCGCCG GCGGTCACAC TACCTATAGA

NeoI

1951 GCAGAACTCT ATCTTCAGGA TCTCGCCATG GAGGGTCTTA GCCTACTCCA
CGTCTTAAGA TACAAGTCCT AGAGCGGTAC CTCCCAGAA CTGATGAGGT

2001 ATTGCCCAGA GATAAATTC GAAAAAGCTC TTTCTTTGTT TGGGTCATCA
TAACGGGTCT CTATTTAAAG CTTTTTCGAG AAAGAAACA ACCCAGTAGT

2051 TCTTATTICA AAAGGCCTTT TCCATGCCTT TGGGTGTTGT GACCAACAGC
AGAATAAAGT TTCCCGGAAA AGGTACGGAA ACCCAACA CA CTGGTTGTCG

2101 ACTTTAGAAG TAACAGAGAT TGACCAGCTA GTCTGCAAGG ATCATCTTGC
TGAAATCTTC ATTGTCTCTA ACTGGTCGAT CAGACGTTCC TAGTAGAACG

2151 ATCAACTGAC CAGCTGAAAT CAGTTGGTCT CAACCTCGAG GGGAGCGGAG
TAGTTGACTG GTCGACTTTA GTCAACCAGA GTTGAGCTC CCCTCGCCTC

EcoRV

2201 TATCTACTGA TATCCCATCT GCGACAAAGC GTTGGGGCTT CAGATCTGGT
ATAGATGACT ATAGGGTAGA CGCTGTTTCG CAACCCCGAA GTCTAGACCA

2251 GTGCCTCCCC AAGTGGTCAG CTATCAAGCA GSAGAATGGG CTGAAAATTG
CACGGAGGGG TTCACCAAGT GATACTTCGT CCTCTTACCC GACTTTTAACT

2301 CTACAATCTT GAAATAAAGA AACCGGACGG GACCGAATGC TTACCCCCAC
GATGTTAGAA CTTTATTTCT TTGGCCTGCC CTCGCTTACG AATGGGGGTG

2351 CGCCGGATGG TGTCAGAGGC TTTCCAAGGT GCGCTATCT TCACAAAGCC
GCGGCCTACC ACAGTCTCCG AAAGGTTCCA CGGCGATACA AGTGTTCGG

2401 CAAGGAACCG GGCCTGCCC GGGTGACTAT GCCTTTCACA AGGATCGGAGC
GTTCTTTGGC CCGGGACGGG CCCACTGATA CGGAAAGTGT TCCTACCTCG

2451 TTTCTTCTC TATGACAGGC TGGCTTCAAC TGTAATTTAC AGAGGAGTCA
AAAGAAGGAG ATACTGTCCG ACCGAAGTTC ACATTAAATG TCTCCTCAGT


```

2501 ATTTTCCTGA CCGGGTAATC GCAITCTTGA TATTGGCTAA ACCAAAGGAA
TAAAACGACT CCCCCATTAG CGTAAGAACT ATAACCGATT TGGTTTCCTT
.....
2551 ACGTTCCCTC AATCACCACC CATTGAGAG GCAGCAAAC ACACAGAAA
TGCAAGGAAG TTAGTGGGGG GTAAGCTCTC CGTCGTTGA TGTGACTTTT
.....
2601 TACATCAAGT TACTATGCCA CATCCTACTT GGAGTACGAA ATCGAAAATT
ATGTAGTTCA ATGATACGGT GTAGCATGAA CCTCATGCTT TAGCTTTTAA
.....
2651 TTGGTGCTCA ACACCTCCAG ACCCTTTTCA AAATTAACAA TAATACTTT
AACCACGAGT TGTGAGGTGC TGGGAAAAGT TTTAATTGTT ATTATGAAA
.....
2701 GTTCTTCTGG ACAGGCCCCA CACGCCTCAG TTCCTTTTCC AGCTGAATGA
CAAGAAGACC TGTCCGGGGT GTGCGGAGTC AAGGAAAAGG TCGACTTACT
.....
2751 TACCATTCAA CTTACCAAC AGTTGAGCAA CACAACGGG AAACATAATT
ATCGTAAGTT GAAGTGCTTG TCAACTCGTT GTGTTGACCC TTTGATTAAA
.....
2801 GGACACTAGA TGCTAATATC AATGCTGATA TTGGTGAATG GGCTTTTGG
CCTGTGATCT ACGATTATAG TTACGACTAT AACCACCTAC CCGAAAAACC
.....
2851 GAAAATAAAA AAATCTCTCC GAACAACAC GTGGAGAAGA GCTGTCTTTC
CTTTTATTTT TTTAGAGAGG CTTGTTGATC CACCTCTTCT CGACAGAAAG
.....
2901 GAACTTTTAT CGCTCAACGA GACAGAAGAC GATGATGCCA CATCGTCGAG
CTTTGAAATA GCGAGTTGCT CTGCTTCTG CTAATACGCT GTAGCAGCTC
.....
2951 AACTACAAGG GGAAGAATCT CCGACCGGGC CACCAGGAAG TATTGCGACC
TTGATGTTTC CCTTCTTAGA GGCTGGCCCG GTGGTCCCTC ATAAGCCTGG
.....
3001 TGGTTCCAAA GGATTCCTCT GGGATGGTTT CATTGCACGT ACCAGAAGGG
ACCAAGGTTT CCTAAGGGGA CCCTACCAA GTAACGTGCA TGGTCTTCCC
.....
3051 GAAACAACAT TGCCGTCTCA GAATTCGACA GAAGGTGCAA GAGTAGATGT
CTTTGTTGTA ACGGCAGAGT CTTAAGCTGT CTTCCAGCTT CTCATCTACA
.....
3101 GAATACTCAG GAAACTATCA CAGAGACAAC TGCAACAATC ATAGGCACTA
CTTATGAGTC CTTTGATAGT GTCTCTGTTG ACGTTGTTAG TATCCGTGAT
.....
3151 ACGGTAACAA CATGCAGATC TCCACCATCG GGACAGGACT GAGCTCCAGC
TGCCATTGTT GTACGTCTAG AGGTGCTAGC CCTGTCTGA CTCGAGGTCT
.....

```

NcoI

```

3201 CAAATCCTGA GTTCCTCACC GACCATGGCA CCAAGCCCTG AACTCAGAC
GTTTAGGACT CAAGGAGTGG CTGGTACCGT GGTTCGGGAC TCTGAGTCTG
.....
3251 CTCCACAACC TACACACCAA AACTACCAGT GATGACCACC GAGGAACCAA
GAGCTGTTGG ATGTGTGCTT TTGATGGTCA CTAATGGTGG CTCCTTGTTT
.....
3301 CAACACCACC GAGAACTCT CCTGGCTCAA CAACAGAAGC ACCCACTCTC
GTGTTGGTGG CTCTTTGAGA GGACCGAGTT GTCTCTTCC TGGGTGAGAG
.....
3351 ACCACCCAG AGAATATAAC AACACGGGT AAAACTGTTT GGGCACAAGA
TGGTGGGGTC TCTTATATTG TTGTCGCCAA TTTTGACAAA CCCGTGTTCT
.....

```

3401 GTCCACAAGC AACGGTCTAA TAACCTCAAC AGTAACACGT ATTCTTGGGA
CAGGTGTTTC TTGCCAGATT ATTGAAGTTG TCATTGTCCA TAAGAACCCT

3451 GCCTTGGACT TCGAAAACGC AGCAGAAGAC AAGTTAACAC CAGGGCCACG
CGGAACCTGA AGCTTTTGCG TCGTCTTCTG TTCAATTGTG GTCCTGGTGC

3501 GGTAAATGCA ATCCCAACTT AACTACTGAG ACTGCACAAG AACAAACATAA
CCATTTACGT TAGGGTTGAA TGTGATGACC TGACGTGTTT TTGTTGTATT

BamHI

3551 TGCTGCTGGG ATTGCCTGGA TCCCGTACTT TGCACCGGGT GCAGAAGGCA
ACGACGACCC TAACGGACCT ACGGCATGAA ACCTGGCCCA CGTCTTCCGT

3601 TATACACTGA AGGCCTTATG CACAACCAA ATGCCTTAGT CTGTGGACTC
ATATGTGACT TCCGGAATAC GTGTTGTTT TACGGAATCA GACACCTGAG

3651 AGACAACCTG CAAATGAAAC AACTCAAGCT CTGCAGCTTT TCTTAAGGGC
TCTGTTGAAC GTTTACTTTG TTGAGTTTCA GACGTCGAAA AGAATTCCTG

3701 CACGACGGAG CTGCGGACAT ATACCATACT CAATAGGAAG GCCATAGATT
GTGCTGCCTC GACGCTGTA TATGGTATGA GTTATCCTTC CGGTATCTAA

BamHI

3751 TCCTTCTGCG ACGATGGGGC GGGACATGTA GSATCCTGGG ACCAGATTGT
AGGAAGACGC TGCTACCCCG CCTGTACAT CCTAGGACCC TGGCTAACA

3801 TGCATTGAGC CACATGATTG GACCAAAAAC ATCACTGATA AAATCAACCA
ACGTAACCTG GTGTACTAAC CTGGTTTTTG TAGTGACTAT TTTAGTTGGT

3851 AATCATCCAT GATTTCATCG ACAACCTTTT ACCCAATCAG GATAATGATG
TTASTAGGTA CTAAGTAGC TGTGGGAAA TGGGTTAGTC CTATTACTAC

BamHI

3901 ATAATTGGTG GACGGGCTGG AGACAGTGGG TCCCTGCAGG AATAGGCATT
TATTAACCAC CTGCCCCGACC TCTGTACCT AGGGACGTCC TTATCCGTAA

3951 ACTGGAATTA TTATTGCAAT CATTGCTCTT CTTTGGCTCT GCAAGCTGCT
TGACCTTAAT AATAACGTTA GTAACGAGAA GAAACGCAGA CGTTCGACGA

BamHI

4001 TGTGGAATA TCAGAAATCC AGCACTGGCG GCCGTTACTA GTGGATCCGA
AACAACTTAT AGTCTTAAGG TCGTGACCGC CGGCAATGAT CACCTAGGCT

NarI

BamHI

XbaI

KasI

BamHI

4051 CCTCGGATCC AAGCTCTAGA CCAGCGGCCT GGATCCAGAT CTGCTGTGCC
CGAGCCTAGG TTGAGATCT GGTCCGCGGA CCTAGGTCTA GACGACACGG

4101 TTCTAGTTC CAGCCATCTG TTGTTTGGCC CTCCCCCGTC CCTTCCTTGA
AAGATCAACC GTCGGTAGAC AACAAACGGG GAGGGGGCAC GGAAGGAACT

4151 CCTTGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT
GGGACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT ACTCCTTTAA

4201 GCATCGCATT GTCTGAGTAG GTGTCATTCT ATTCTGGGGG GTGGGGTGGG
CGTAGCGTAA CAGACTCATC CACAGTAAGA TAAGACCCCC CACCCACCCC

SphI

4251 CCAGCACAGC AAGCGGGAGG ATTGGGAAGA CAATAGCAGG CATGCTGGGG
CGTCGTGTCT TCCCCCTCC TAACCCTTCT GTTATCGTCC GTACGACCCC

KpnI

4301 ATGCGGTGGG CTCTATGGGT ACCCAGGTGC TGAAGAATTG ACCCGGTTC
TACGCCACCC GAGATACCCA TGGGTCCACG ACTTCTTAAC TGGGCCAAGG

4351 TCCTGGGCCA GAAAGAAGCA GGCACATCCC CTTCTCTGTG ACACACCCTG
AGGACCCCGT CTTTCTTCGT CCGTGTAGGG GAAGAGACAC TGTGTGGGAC

4401 TCCAGCCCC TGGTTCTTAG TTCCAGCCCC ACTCATAGGA CACTCATAGC
AGGTGCGGGG ACCAAGAATC AAGGTGCGGG TGAGTATCCT GTGAGTATCG

4451 TCAGGAGGGC TCCGCCCTCA ATCCACCCCG CTAAAGTACT TGGAGCGGTC
AGTCCCTCCC AGCGGGAAGT TAGGGTGGGC GATTTTCATGA ACCTCGCCAG

4501 TCTCCCTCCC TCATCAGCCC ACCAAACCAA ACCTAGCCTC CAAGAGTGGG
AGAGGGAGGG AGTAGTCGGG TGGTTGGTT TCGATCGGAG GTTCTCACCC

4551 AAGAAATTAA AGCAAGTAG GCTATTAAGT GCAGAGGGAG AGAAATGCC
TTCTTTAATT TCGTTCTATC CGATAATTCA CGTCTCCCTC TCTTTACGG

XmnI

4601 TCCAACATGT GAGGAAGTAA TCAGAGAAAT CATAGAATTT CTTCGGCTTC
AGGTTGTACA CTCCTTCATT ACTCTCTTA GTATCTTAA GAAGGCCAAG

4651 CTCGCTCACT GACTCGCTGC GCTCGGTCTG TCGGCTCCGG CGAGCGGTAT
GAGCGAGTGA CTGAGCGAGG CGAGCCAGCA AGCCGACGCC CCTCGCCATA

4701 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG

4751 GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
CGTCCTTTCT TGTACACTCG TTTTCCGGTC GTTTTCCGGT CCTTGGCATT

4801 AAAGCCCGCG TTCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC
TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG

4851 ATCACAATAA TCGACGCTCA ACTCAGAGGT GGCGAAACCC GACAGGACTA
TAGTGTTTTT AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGAT

4901 TAAAGATACC AGCGGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
ATTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAGC CGAGACGACA

4951 TCCGACCCTG CCGCTTACCG GATACCTCTC CGCCTTCTC CTTTCGGGAA
AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT

5001 GCGTGGCGCT TTCTCAATCC TCACGCTGTA GGTATCTCAG TTCGGTGTAG
CGCACCGCCA AAGAGTTACG AGTGCGACAT CCATAGACTC AAGCCACATC

```

5051  GTCGTTCCGT  CCAAGCTGGG  CTGTGTGCAC  GAACCCCCCG  TTCAGCCCGA
      CAGCAAGCGA  GGTTCGACCC  GACACACGTG  CTTGGGGGGC  AAGTCGGGCT
.....
5101  CCGCTGCCCC  TTATCCGGTA  ACTATCGTCT  TGAGTCCAAC  CCGGTAAAGAC
      GCGGACGCGG  AATAGGCCAT  TGATAGCAGA  ACTCAGGTTG  GGCCATTCTG
.....
5151  ACGACTTATC  GCCACTGGCA  GCAGCCACTG  GTAACAGGAT  TAGCAGAGCG
      TGCTGAATAG  CGGTGACCGT  CGTCGGTGAC  CATGTCTCTA  ATCGTCTCGC
.....
5201  AGGTATGTAG  GCGGTGCTAC  AGAGTCTTGG  AAGTGGTGGC  CTAACTACGG
      TCCATACATC  CGCCACGATG  TCTCAAGAAC  TTCACCACCG  GATTGATGCC
.....
5251  CTACACTAGA  AGGACAGTAT  TTGGTATCTG  CGCTCTGCTG  AAGCCAGTTA
      GATCTGATCT  TCCTGTCTAT  AACCATAGAC  GCGAGACGAC  TTCGGTCAAT
.....
5301  CCTTCGGAAG  AAGAGTTGGT  AGCTCTTGAT  CCGGCAAAAC  AACCACCGCT
      GGAAGCCTTT  TTCTCAACCA  TCGAGAACTA  GGCCGTTTGT  TTGGTGGCGA
.....
5351  GGTAGCGGTG  GTTTTTTTGT  TTGCAAGCAG  CAGATTACGC  GCAGAAAAAA
      CCATCGCCAC  CAAAAAAACA  AACGTTCTGT  GTCTAATGCG  CGTCTTTTTT
.....
5401  AGGATCTCAA  GAAGATCCTT  TGATCTTTTC  TACGGGGTCT  GACGCTCAGT
      TCCTAGAGTT  CTTCTAGGAA  ACTAGAAAAG  ATGCCCCAGA  CTGCGAGTCA
.....
5451  GGAACGAAAA  CTCACGTTAA  GGGATTTTGG  TCATGAGATT  ATCAAAAAGG
      CCTTGCTTTT  GAGTGCAATT  CCCTAAAACC  AGTACTCTAA  TAGTTTTTCC
.....
5501  ATCTTCACCT  AGATCCTTTT  AAATTAAAAA  TGAAGTTTTA  AATCAATCTA
      TAGAAGTGGA  TCTAGGAAAA  TTTAATTTTT  ACTTCAAAAT  TTAGTCAAGT
.....
5551  AAGTATATAT  GAGTAAACTT  GGTCTGACAG  TTACCAATGC  TTAATCAGTG
      TTCATATATA  CTCATTTGAA  CCAGACTGTC  AATGGTTACG  AATTAGTCAC
.....
5601  AGGCACCTAT  CTCAGCGATC  TGTCTATTTT  GTTCATCCAT  AGTTCCTTGA
      TCCGTGGATA  CAGTCGCTAG  ACAGATAAAG  CAAGTAGGTA  TCAACGGACT
.....
5651  CTCCGGGGGG  GGGGGGGCGT  GAGGTCTGCC  TCGTGAAGAA  GGTGTTGCTG
      GAGGCCCCCC  CCCCCCGCGA  CTCCAGACGG  AGCACTTCTT  CCACAACGAC
.....
5701  ACTCATACCA  GGCCTGAATC  GCCCCATCAT  CCAGCCAGAA  AGTGAGGGAG
      TGAGTATGGT  CCGGACTTAG  CGGGGTAGTA  GGTCGGTCTT  TCACTCCCTC
.....
5751  CCACGGTTGA  TGAGAGCTTT  GTTGTAGGTG  GACCAGTTGG  TGATTTTGAA
      GGTGCCAACT  ACTCTCGAAA  CAACATCCAC  CTGGTCAACC  ACTAAAACCT
.....
5801  CTTTTCCTTT  GCCACGGAAC  GGTCTGCGTT  GTCCGGAAGA  TCCCTGATCT
      GAAAACGAAA  CGGTGCCCTG  CCAGACGCAA  CAGCCCTTCT  ACGCACTAGA
.....
5851  GATCCTTCAA  CTCAGCAAAA  GTTCGATTTA  TTCAACAAAG  CCGCCGTCCC
      CTAGGAAGTT  GAGTCGTTTT  CAAGCTAAAT  AAGTTGTTTC  GCGGCGAGGG
.....
5901  GTCAAGTCAG  CGTAATGCTC  TGCCAGTGT  ACAACCAATT  AACCAATTCT
      CAGTTCAGTC  GCATTACGAG  ACGGTCACAA  TGTGTTTAA  TTGTTAAGA
.....
5951  GATTAGAAAA  ACTCATCGAG  CATCAATGA  AACTGCAATT  TATTCATATC
      CTAATCTTTT  TGAGTAGCTC  GTAGTTTACT  TTGACGTTAA  ATAAGTATAG
.....

```

6001 AGGATTATCA ATACCATATT TTTGAAAAAG CCGTTTCTGT AATGAAGGAG
TCCTAATAGT TATGGTATAA AAACCTTTTC GCCAAAGACA TTACTTCCTC

6051 AAAACTCACC GAGGCAGTTC CATAGCATGG CAAGATCCTG GTATCGGTCT
TTTTGAGTGG CTCGCTCAAG GTATCCTACC GTTCTAGGAC CATAGCCAGA

6101 GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAAAT TCCCCTCGTC
CGCTAAGGCT GAGCAGGTTG TAGTTATGTT CGATAATTAA AGGGGACCAG

6151 AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGACG ACTGAATCCG
TTTTTATTCC AATAGTTCAC TCTTTAGTGG TACTCACTGC TGACTTAGGC

HindIII

6201 GTGAGAATGG CAAAAGCTTA TGCATTCTT TCCAGACTTG TTCAACAGGC
CACTCTTACC GTTTTCGAAT ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG

6251 CAGCCATTAC GCTCGTCATC AAAATCACTC GCATCAACCA AACCGTTATT
GTCCGTAATG CGAGCAGTAG TTTTAGTGAG CGTAGTTGGT TTGGCAATAA

PvuI

6301 CATTCTGTAT TCGGCCTGAG CGAGACGAAA TACGCGATCG CTCTTAAAG
GTAAGCACTA ACGCGGACTC GCTCTGCTTT ATGCGCTAGC GACAATTTTC

6351 GACAATTACA AACAGGAATC GAATGCAACC GCGCCAGGAA CACTGCCAGC
CTGTTAATGT TTGTCTTAG CTACGTTGG CCGCGTCTT GTGACGGTCG

6401 GCATCAACAA TATTTTCACC TGAATCAGGA TACTCTTCTA ATACCTGGAA
CGTAGTTGTT ATAAAGTGG ACTTAGTCCT ATAAGAAGAT TATGGACCTT

6451 TGCTGTTTTT CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG
ACGACAAAAG GGCCCTAGC GTCACCACTC ATTGGTACGT AGTAGTCCTC

6501 TACGGATAAA ATGCTTGATG GTCGGAAGAG GCATAAATTC CGTCAGCCAG
ATGCCTATTT TACGAACTAC CAGCCTTCTC CGTATTTAAG GCAGTCGGTC

6551 TTAGTCTGA CCATCTCATC TGTAACTCA TTGGCAACGC TACCTTTGCC
AAATCAGACT GGTAGAGTAG ACATTGTAGT AACCGTTGCG ATGGAAACGG

ClaI

6601 ATGTTTCAGA AACAACTCTG GCGCATCGGG CTTCCTATAC AATCGATAGA
TACAAAGTCT TTGTTGAGAC CGCGTAGCCC GAAGGGTATG TTAGCTATCT

6651 TTGTCCGACC TGATTGCCCG ACATTATCGC GAGCCCATTT ATACCCATAT
AACAGCGTGG ACTAACGGGC TGTAAATAGCG CTCGGGTAAA TATGGGTATA

6701 AAATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AAGACGTTTC
TTAGTCGTA GGTACAACCT TAAATTAGCG CCGGAGCTCG TTCTGCAAG

6751 CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG
GGCAACTTAT ACCGAGTATT GTGGGGAACA TAATGACAAA TACATTCTGC

6801 ACAGTTTTAT TGTTTCATGAT GATATATTTT TATCTTGTGC AATGTAACAT
TGTCAAAATA ACAAGTACTA CTATATAAAA ATAGAACACG TTACATTGTA

DraIII

```
6851 CAGAGATTTT GAGACACAAC GTGGCTTTCC CCCCCCCCCC ATTATTGAAG
      CTCTCTAAAA CTCTGTGTTC CACCGAAAGG GGGGGGGGGC TAATAACTTC
.....
6901 CATTATCAG GGTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT
      GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATRAA CTTACATAAA
.....
6951 AGAAAAATAA ACAAAATAGG GTTCCCGCCA CATTCCCCG AAAAGTGCCA
      TCTTTTTATT TGTTTATCCC CAAGGCGCGT GTAAAGGGGC TTTTCACGCT
.....
7001 CCTGACGTCT AAGAAACCAT TATTATCATG ACATTACCT ATAAAAATAG
      GGAATGCAGA TTCTTTGGTA ATAATAGTAC TGAATTGGA TATTTTATC
.....
7051 GCGTATCAG AGGCCCTTTC GTC
      CGCATAGTGC TCCGGGAAAG CAG
.....
```

pVR 1012-GP(Z)

General Description

DNA pVR 1012-GP(Z)
 Local object
 Created: 09/15/98 05:06PM
 Last Modification Date: ? (no data)
 length: 7285 bp
 storage type: Basic
 form: Circular

Comments

Sequence Listing ID No: 3

Restriction Map

DraIII: 1 site CACNANGTG
 GTGNNACAC

HindIII: 1 site AAGCTT
 TTCGAA

HpaI: 1 site GTTAAC
 CAATTG

KasI: 1 site GGCGCC
 CCGCGG

NarI: 1 site GGCGCC
 CCGCGG

NotI: 1 site GCGGCCGC
 CGCCGCGC

PmlI: 1 site CACGTG
 GTGCAC

PvuI: 1 site CGATCG
 GCTAGC

SacII: 1 site CCGCGG
 GCGGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATGG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATGC
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1860

TbGH

Start: 4302 End: 4854

Kan r

Start: 6350 End: 6972 (Complementary)

Misc_f ature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(Z)

Start: 1870 End: 4301

Annotations


```

1  TCGCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
   AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC
.....
51  GAGACGGTCA CAGCTTGTCT GTAACCGGAT CCGGGGAGCA GACAAGCCCC
   CTCTGCCAGT GTCGAACAGA CATTTCGCTA CGGCCCTCGT CTGTTGCGGC
.....
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACC GCCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
   CCGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACGC CACACTTTAT
.....
201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
   GCGGTGTCTA CCGATTCCCT TTTTATGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
   AACGTATCCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGGTTGTAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
   TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA
.....
401 ACATAACTTA CCGTAATGCG CCCGCTGGC TGACCGCCCA ACGACCCCG
   TGTATTGAAT CCCATTACCG GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
   GGGTAAGTGC AGTTATTACT GCATACAAGG GTATCATTCG GGTATCCCT
.....
501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCATGTAG TTCACATAGT ATACGGTTCA TCGGGGGGAT AACTGCAGTT
.....
601 TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTATGTATC TGAATATCCC
.....

```

NcoI

```

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TGAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
   CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACGTAG
.....
751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCCAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA
.....
801 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
   CCGTGCTTTT AGTTGCCCTG AAAGGTTTAA CAGCATTGTT GAGGCGGGGT
.....

```

851 TTGACGCAAA TGGGCGGTAG GCGTGACGG TCGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAAT CACTTGCCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG CCCCTGGCTA GGTGCGAGGC GCCGGCCCTT

1001 CCGTGCAATTG GAACCGCGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGACGCG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GACATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG CGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGCGGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
TGCCTGACAC ATAAAAATGT CCTACCCAG GGTAAATAAT AAATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTC ATTAAACATA
GTGTATATGT TGTTGCGGCA GGGGGCACGG CCGTCAAAAA TAATTGTAT

1401 GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
CGCACCCATG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
AGAGGCCATC CCCGCCCTGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCCG CAGCTCCTTG CTCCTAACAG TGGAGCCAG
TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAAATG CACCCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGCAGATTG GGCTCGCACG
ACCCCATCC CATAACAGA CTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GGTGACGCAG ATGGAAGACT TAAGCCAGCG GCAGAAGAAG ATGCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGCT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
GACTCAACAA CATAAGACTA TTCTCACTCT CCATTGAGGG CAACGCCACG

HpaI
 1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
 ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI
 1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
 GCGCGGTGGT CTGTATTATC GACTGTCTCA TTGTCTGACA AGGAAAGGTA

NcoI PmlI EcoRVNotI
 1851 GGGTCTTTTC TGCACTCACC GTCCTCGACA CGTGTGATCA GATATCGCGG
 CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

NarI
NotI XbaI KasI
 1901 CCGCTCTAGA CCAGGCGCCT GGATCGATCC GCGATGAAGA TTAAGCCGAC
 GCGGAGATCT GGTCCGCGGA CCTAGCTAGG CGCTACTTCT AATTCGGCTG

1951 AGTGAGCGTA ATCTTCATCT CTCTTAGATT ATTTGTTTTC CAGAGTAGGG
 TCACTCGCAT TAGAAGTAGA GAGAATCTAA TAAACAAAAG GTCTCATCCC

2001 GTCGTCAGGT CCTTTTCAAT CGTGTAACCA AAATAAACTC CACTAGAAGG
 CAGCAGTCCA GGAAAAGTTA GCACATTGGT TTTATTGAG GTGATCTTCC

2051 ATATTGTGGG GCAACAACAC AATGGGCGTT ACAGGAATAT TCCAGTTACC
 TATAACACCC CGTTGTTGTG TTACCCGCAA TGTCCTTATA ACGTCAATGG

2101 TCGTGATCGA TTCAAGAGGA CATCATCTCT TCTTTGGGTA ATTATCCTTT
 AGCACTAGCT AAGTTCCTCT GTAGTAAGAA AGAAACCCAT TAATAGGAAA

2151 TCCAAAGAAC ATTTTCCATC CCACTTGGAG TCATCCACAA TAGCACATTA
 AGGTTTCTTG TAAAGGTAG GGTGAACCTC AGTAGGTGTT ATCGTGTAAT

2201 CAGGTTAGTG ATGTGCACAA ACTAOTTTGT CGTGACAAAC TGTCATCCAC
 GTCCAATCAC TACAGCTGTT TGATCAAACA GCACTGTTTG ACAGTAGGTG

2251 AAATCAATTG AGATCAGTTG GACTGAATCT CGAAGGGAAT GGAGTGGCAA
 TTTAGTTAAC TCTAGTCAAC CTGACTTAGA GCTTCCCTTA CCTCACCCTT

2301 CTGACGTGCC ATCTGCAACT AAAAGATGGG GCTTCAGGTC CGGTGTCCCA
 GACTGCACGG TAGACGTTGA TTTTCTACCC CGAAGTCCAG GCCACAGGGT

2351 CCAAAGGTGG TCAATTATGA AGCTGGTGAA TGGGCTGAAA ACTGCTACAA
 GGTTCACACC AGTTAATACT TCGACCACTT ACCCGACTTT TGACGATGTT

2401 TCTTGAAATC AAAAAACCTG ACGGCAGTGA GTGTCTACCA GCAGCGCCAG
 AGAAGTTTAG TTTTGTGGAC TGCCCTCACT CACAGATCGT CGTCGCGGTC

2451 ACGGGATTCG GGGCTTCCCC CGGTGCCCGT ATGTGCACAA AGTATCAGGA
 TGCCCTAACC CCCGAAGGGG GCCACGGCCA TACACGTGTT TCATAGTCCT

2501 ACGGGACCGT GTGCCGAGA CTTTGCCCTC CATAAAGAGG GTGCTTCTTT
 TGCCCTGGCA CACGGCCTCT GAAACGGAAG GTATTTCTCC CACGAAAGAA

2551 CCTGTATGAT CGACTTGCTT CCACAGTTAT CTACCGAGGA ACGACTTTCC
GGACATCTA GCTGAACGAA GGTGTCAATA GATGGCTCCT TGCTGAAAGC

2601 CTGAAGGTGT CGTTGCATTT CTGATACTGC CCCAAGCTAA GAAGGACTTC
GACTTCCACA GCAACGTAAA GACTATGACG GGGTTCGATT CTTCCTGAAG

2651 TTCAGCTCAC ACCCCTTGAG AGAGCCGGTC AATGCAACGG AGGACCCGTC
AAGTCGAGTG TGGGAATC TCTCGGCCAG TTACGTTGCC TCCTGGGCAG

EcoRV

2701 TAGTGGCTAC TATTCTACCA CAATTAGATA TCAGGCTACC GGTTTTGGAA
ATCACCAGATG ATAAGATGGT GTTAATCTAT AGTCCGATGG CCAAAACCTT

2751 CCAATGAGAC AGAGTACTTG TTCGAGGTTG ACAATTTGAC CTACGTCCAA
GGTACTCTG TCTCATGAAC AAGCTCCAAC TGTTAAACTG GATGCAGGTT

2801 CTTGAATCAA GATTACACAC ACAGTTTCTG CTCCAGCTGA ATGAGACAAT
GAACCTAGTT CTAAGTGTGG TGTCAAAGAC GAGGTCGACT TACTCTGTTA

2851 ATATACAAGT GGGAAAAGCA GCAATACCAC GGGAAAAC TAATTGGAAGG
TATATGTTCA CCCTTTTCCT CGTTATGGTG CCCTTTTGAT TAAACCTTCC

2901 TCAACCCCGA AATTGATACA ACAATCGGGG AGTGGGCTT CTGGGAAACT
AGTTGGGGCT TTAACATATGT TGTTAGCCCC TCACCCGGAA GACCCCTTGA

2951 AAAAAAACC TCACTAGAAA AATTGCGAGT GAAGAGTTGT CTTTACAGT
TTTTTTTTGG AGTGATCTTT TTAAGCGTCA CTTCTCAACA GAAAGTGTC

3001 TGTATCAAAC GGAGCCAAAA ACATCAGTGG TCAGAGTCCG GCGCGAACTT
ACATAGTTTG CCTCGGTTTT TGTAGTCACC AGTCTCAGGC CGCGCTTGAA

3051 CTTCCGACCC AGGGACCAAC ACAACAAC TGAGACCACAA AATCATGGCT
GAAGGCTGGC TCCCTGGTTG TGTGTTGAC TTCTGGTGT TTAGTACCGA

3101 TCAGAAAATT CCTCTGCAAT GGTTCAGTG CACAGTCAAG GAAGCGAAGC
AGTCTTTTAA GGAGACGTTA CCAAGTTCAC GTGTCAGTTC CTTCCCTTCG

3151 TCCAGTGTCTG CATCTAACAA CCCTTGCCAC AATCTCCAGC AGTCCCAAT
ACGTCACAGC GTAGATTGTT GGAACGGTG TTAGAGGTGC TCAGGGGTTA

3201 CCCTCACAA CAAACCAGGT CCGGACAACA GCACCCATAA TACACCCGTG
GGGAGTGTG GTTTGGTCCA GGCTGTGTGT CGTGGGTATT ATCTGGGCAC

3251 TATAAACTTG ACATCTCTGA GCGAACTCAA GTTGAACAAC ATCACCAGCAG
ATATTGGAAC TGTAAGAGCT CCGTTCAGTT CAACCTGTTG TAGTGGCGTC

3301 AACAGACAC GACAGCACAG CCTCCGACAC TCCCTCTGCC ACGACCCGAG
TTGTCTGTTG CTCTCGTGTC GGAGGCTGTG AGGGAGACGG TGCTGGCGTC

3351 CCGGACCCCC AAAAGCAGAG AACACCAACA CGAGCAAGAG CACTGACTTC
GGCCTGGGGG TTTTCGTCTC TTGTGGTTGT GCTCGTTCTC GTGACTGAAG

3401 CTGGACCCCG CCACCACAAC AAGTCCCCAA AACCACAGCG AGACCGCTGG
GACCTGGGCG GGTGGTGTG TTCAGGGGTT TTGGTGTGCG TCTGGCGACC

```

3451 CAACAACAAC ACTCATCACC AAGATACCGG AGAAGAGAGT GCCAGCAGCG
    GTTGTGTTG TGAGTAGTGG TTCTATGGCC TCTTCTCTCA CGGTGCTCGC
.....
3501 GGAAGCTAGG CTTAATTACC AATACTATTG CTGGAGTCGC ACGACTGATC
    CCTTCGATCC GAATTAATGG TTATGATAAC GACCTCAGCG TCCTGACTAG
.....
3551 ACAGGCGGGA GAAGAACTCG AAGAGAAGCA ATTGTCAATG CTCAACCCAA
    TGTCGCGCCT CTTCTTGAGC TTCTCTTCGT TAACAGTTAC GAGTTGGGTT
.....
3601 ATCCAACCCCT AATTTACATT ACTGGACTAC TCAGGATGAA GGTGCTGCAA
    TACGTTGGGA TTAATGTAA TGACCTGATG AGTCCTACTT CCACGACGTT
.....
3651 TCGGACTGGC CTGGATACCA TATTTCGGGC CAGCAGCCGA GGAATTTAC
    AGCCTGACCG GACCTATGGT ATAAAGCCCG GTCGTGCGCT CCCTTAAATG
.....
3701 ATAGAGGGCC TAATGCACAA TCAAGATGGT TTAATCTGTG GGTGAGACA
    TATCTCCCGG ATTACGTGTT AGTTCTACCA AATTAGACAC CCAACTCTGT
.....
3751 GCTGGCCAC CAGACGACTC AAGCTCTTCA ACTGTTCCCTG AGAGCCACAA
    CGACCGGTTG CTCTGCTGAG TTCGAGAAGT TGACAAGGAC TCTCGGTGTT
.....
3801 CTGAGCTACG CACCTTTTCA ATCCTCAACC GTAAGGCAAT TGATTCTTG
    GACTCGATGC GTGGAAAAGT TAGGAGTTGG CATTCCTTAA ACTAAAGAAC
.....
3851 CTGCAGCGAT GGGGCGGCAC ATGCCACATT CTGGGACCGG ACTGCTGTAT
    GACGTCGCTA CCCC GCCGTG TACGGTGTA GACCCTGGCC TGACGACATA
.....
3901 CGAACCACAT GATTGGACCA AGAACATAAC AGACAAAAT GATCAGATTA
    GCTTGGTGTA CTAACCTGGT TCTTGTAATTG TCTGTTTTAA CTAGTCTAAT
.....
3951 TTCATGATT TGTTGATAAA ACCCTTCCGG ACCAGGGGGA CAATGACAA
    AAGTACTAAA ACAACTATTT TGGGAAGGCC TGGTCCCCCT GTTACTGTTA
.....
4001 TGGTGGACAG GATGGAGACA ATGGATACCG GCAGGTATTG GAGTTACAGG
    ACCACCTGTC CTACCTCTGT TACCTATGGC CGTCCATAAC CTCAATGTCC
.....
4051 CGTTATAATT GCAGTTATCG CTTTATTCTG TATATGCAA TTTGCTTTT
    GCAATATTAA CGTCAATAGC GAAATAAGAC ATATACGTTT AAACAGAAAA
.....
4101 AGTTTTTCTT CAGATTGCTT CATGCAAAAG CTCAGCCTCA AATCAATGAA
    TCAAAAACAA GTCTAACGAA GTACCTTTTC GAGTCGGAGT TTAGTTACTT
.....
4151 ACCAGGATTT AATTATATGG ATTACTTGAA TCTAAGATTA CTGACAAAT
    TGGTCCTAAA TTAATATACC TAATGAACCT AGATTCTAAT GAAGCTGTTA
.....
4201 GATAATATAA TACACTGGAG CTTTAAACAT AGCCAATGTG ATTCTAACTC
    CTATTATATT ATGTGACCTC GAAATTTGTA TCGGTTACAC TAAGATTGAG
.....
4251 CTTTAAACTC ACAGTTAATC ATAAACAAGG TTTGGTACCG AGCTCGAATT
    GAAATTTGAG TGCAATTAG TATTTGTTCC AAACCATGGC TCGAGCTTAA
.....
4301 ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTGTTTGC CCCTCCCCCG
    TAGACGACAC GGAAGATCAA CGGTCGGTAG ACAACAAACG GGGAGGGGGC
.....
4351 TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
    ACGGAACGAA CTGGGACCTT CCACGGTGAG GGTGACAGGA AAGGATTATT
.....

```

4401 AATGAGGAAA TTGCATCCCA TTGTCTGAGT AGGTGTCAAT CTATTCTGGG
TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC

4451 GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA GACAATAGCA
CCCACCCAC CCCGTCGTGT CGTTCCTCCCT CCTAACCTT CTGTATCGT

SphI

4501 GGCATGCTCG GGATCGGGTG GGCTCTATGG GTACCCAGGT GCTGAAGAAT
CCGTACGACC CCTACGCCAC CCGAGATACC CATGGGTCCA CGACTTCTTA

4551 TGACCCGGTT CCTCTGGGC CAGAAAGAAG CAGGCACATC CCCTTCTCTG
ACTGGGCCAA GGAGGACCCG GTCTTCTTC GTCCGTGTAG GGAAGAGAC

4601 TGACACACCC TGTCCAGCC CCTGGTTCTT AGTTCAGCC CCACTCATAG
ACTGTGTGGG ACABGTGCGG GGACCAAGAA TCAAGGTGCG GTGAGTATC

4651 GACACTATA GCTCAGGAGG GCTCCGCTT CAATCCCACC CGCTAAAGTA
CTGTGAGTAT CGAGTCTCC CGAGGCGGAA GTTAGGGTGG CCGATTTCAT

4701 CTGGAGCGG TCTCTCCCTC CCTCATCAGC CCACCAACC AAACCTAGCC
GAACCTCGCC AGAGAGGGAG CGAGTAGTCC GGTGGTTGG TTTGGATCGG

4751 TCCAAGAGTG GGAAGAAATT AAAGCAAGAT AGGCTATTAA GTGCAGAGGG
AGGTTCTCAC CCTTCTTTAA TTCTGTTCTA TCCGATAATT CACGTCTCCC

4801 AGAGAAATG CCTCCAACAT GTGAGGAAGT AATGAGAGAA ATCATAGAAT
TCTCTTTTAC GGAGGTGTGA CACTCCTTCA TTAATCTCTT TAGTATCTTA

4851 TTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
AAGAAGCGCA AGGAGCGAGT GACTGAGCGA CCGAGCCAG CAAGCCGACG

4901 GCGAGCGGT ATCAGCTCAC TCAAGGCGG TAATACGGT ATCCACAGAA
CCGCTCGCCA TAGTCGAGTG AGTTCCGCG ATTATGCCAA TAGGTGTCTT

4951 TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGCCC AGCAAAAGGC
AGTCCCTAT TCGTCTCTT CTGTGACACT CGTTTTCCGG TCGTTTTCCG

5001 CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
GTCCTTGGCA TTTTTCCGGC GCAACGACCG CAAAAGGTA TCCGAGGCGG

5051 CCCCTGACGA GCATACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC
GGGACTGCT CGTAGTGTGTT TAGCTCGGA GTTCAGTCTC CACCGCTTG

5101 CCGACAGGAC TATAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
GGCTGTCTG ATATTTCTAT GGTCCGCAAA GCGGGACCTT CGAGGGAGCA

5151 GCGCTCTCCT GTTCCGACCC TCCCGCTTAC CGGATACCTG TCCGCCTTTC
CGCGAGAGGA CAAGGCTGGC ACGGCGAATG GCCTATGGAC AGCGGAAAG

5201 TCCCTTCGGG AAGCGTGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
AGGGAAGCCC TTCGCACCGC GAAAGAGTTA CGAGTGCGAC ATCCATAGAG

5251 AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
TCAAGCCACA TCCAGCAAGC GAGGTTCCAC CCGACACAG TGCTTGGGGG

```

5301  CGTTCAGCCC  GACCGCTGCG  CCTTATCCCG  TAACTATCGT  CTTGAGTCCA
      GCAAGTCGGG  CTGGCGACGC  GGAATAGGCC  ATTGATAGCA  GAACTCAGGT
.....
5351  ACCCGGTAAG  ACACGACTTA  TCGCCACTGG  CAGCAGCCAC  TGGTAACAGG
      TGGGCCATT  TGTGCTGAAT  AGCGGTGACC  GTCGTGGGTG  ACCATTGTCC
.....
5401  ATTAGCAGAG  CGAGGTATGT  AGCCGGTGCT  ACAGAGTTCT  TGAAGTGGTG
      TAATCGTCTC  GCTCCATACA  TCCGCCACGA  TGTCTCAAGA  ACTTCACCAC
.....
5451  GCCTAACTAC  GGCTACACTA  GAAGGACAGT  ATTTGGTATC  TGGCCTCTGC
      CGGATTGATG  CCGATGTGAT  CTTCTGTCTA  TAAACCATAG  ACCCGAGACG
.....
5501  TGAAGCCAGT  TACCTTCGGA  AAAAGAGTTG  GTAGCTCTTG  ATCCGGCAAA
      ACTTCGGTCA  ATGGAAGCCT  TTTTCTCAAC  CATCGAGAAC  TAGGCCGTTT
.....
5551  CAAACCCACG  CTGGTAGCGG  TGGTTTTTTT  GTTTGCAAGC  AGCAGATTAC
      GTTTGGTGGC  GACCATCGCC  ACCAAAAAAA  CAAACGTTCC  TCGTCTAATG
.....
5601  CCGCAGAAAA  AAAGGATCTC  AAGAAGATCC  TTTGATCTTT  TCTACGGGGT
      CGCCTCTTTT  TTTCTTAGAG  TTCTTCTAGG  AACTAGAAAA  AGATGCCCCA
.....
5651  CTGACCGTCA  GTGGAACGAA  AACTCAGGTT  AAGGGATTTT  GGTGATGAGA
      GACTGCGAGT  CACCTTGCTT  TTGAGTGCAA  TTCCCTAAAA  CCAGTACTCT
.....
5701  TTATCAAAAA  GGATCTTCAC  CTAGATCCTT  TTAAATTAAA  AATGAAGTTT
      AATAGTTTTT  CCTAGAAGTG  GATCTAGCAA  AATTTAATTT  TTAATTCAAA
.....
5751  TAAATCAATC  TAAAGTATAT  ATGAGTAAAC  TTGGTCTGAC  AGTTACCAAT
      ATTTAGTTAG  ATTTATATA  TACTCATTG  AACCAGACTG  TCAATGGTTA
.....
5801  CCTTAATCAG  TGAGGCACCT  ATCTCAGCGA  TCTGTCTATT  TCCTTCATCC
      CGAATTAGTC  ACTCCGTGGA  TAGAGTCGCT  AGACAGATAA  AGCAAGTAGG
.....
5851  ATAGTTGCCT  GACTCCGCGG  GGGGGGGGCG  CTGAGGTCTG  CCTCGTGAAG
      TATCAACCGA  CTGAGGCCCC  CCCCCCCCCG  GACTCCAGAC  GGAGCACTTC
.....
5901  AAGGTGTTGC  TGACTCATA  CAGGCCTGAA  TCGCCCCATC  ATCCAGCCAG
      TTCCACAACG  ACTGAGTATG  GTCCGGACTT  AGCGGGGTAG  TAGGTGCGTC
.....
5951  AAAGTGAGGG  AGCCACGGTT  GATGAGAGCT  TTGTTGTAGG  TGGACCAGTT
      TTTCACTCCC  TCGGTCCCAA  CTACTCTCGA  AACAACATCC  ACCTGGTCAA
.....
6001  GGTGATTTTG  AACTTTTGCT  TTGCCACGGA  ACGGTCTGCG  TTGTCGGGAA
      CCACTAAAC  TTGAAAACGA  AACGGTGCC  TGCCAGACGC  AACAGCCCTT
.....
6051  GATGCGTGAT  CTGATCCTTC  AACTCAGCAA  AAGTTCGATT  TATTCAACAA
      CTACGCACTA  GACTAGGAAG  TTGAGTCGTT  TTCAAGCTAA  ATAAGTTGTT
.....
6101  AGCCGCCGTC  CCGTCAAGTC  AGCGTAATGC  TCTGCCAGTG  TTACAACCAA
      TCGCGCGCAG  GGCAGTTCAG  TCGCATTACG  AGACGGTCAC  AATGTTGGTT
.....
6151  TTAACCAATT  CTGATTAGAA  AAACATCATG  AGCATCAAAT  GAAACTGCAA
      AATGGTTAA  GACTAATCTT  TTTGAGTAGC  TCGTAGTTTA  CTTTGACGTT
.....
6201  TTTATTCATA  TCAGGATTAT  CAATACCATA  TTTTGA AAAA  AGCCGTTTCT
      AAATAAGTAT  AGTCTAATA  GTTATGGTAT  AAAAATTTT  TCGGCAAGA
.....

```

6251 GTAATGAAGG AGAAAACCTCA CCGAGGCAGT TCCATAGGAT GGCAAGATCC
CATTACTTCC TCTTTTGAGT GGCTCCCTCA AGGTATCCTA CCGTTCTAGG

6301 TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
ACCATAGCCA GACGCTAAGG CTGACCAGGT TGTAAGTTATG TTGGATAATT

6351 TTTCCCCTCG TCAAAAATAA GGTATCAAG TGAGAAATCA CCATGAGTGA
AAAGGGGAGC AGTTTTTATT CCAATAGTTC ACTCTTTAGT GGTACTCACT

HindIII

6401 CGACTGAATC CGGTGAGAAT GGCAAAAGCT TATGCATTTC TTTCCAGACT
GCTGACTTAG GCCACTCTTA CCGTTTTCGA ATACGTAAAG AAAGGTCTGA

6451 TGTTC AACAG GCCAGCCATT ACGCTCGTCA TCAAAATCAC TCGCATCAAC
ACAAGTTGTC CGGTCGGTAA TSCGAGCAGT AGTTTTAGTG AGCGTAGTTG

PvuI

6501 CAAACCGTTA TTCATTCGTG ATTGCCGCTG AGCGAGACGA AATACGCGAT
GTTTGGCAAT AAGTAAGCAC TAACCGGAC TCGCTCTGCT TTATGCCGTA

PvuI

6551 CCGTGTTAAA AGGACAATTA CAAACAGGAA TCGAATGCAA CCGGCGCAGG
CGGACAATTT TCCTGTTAAT GTTTGTCCTT AGCTTACGTT GGCCGCGTCC

6601 AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
TTGTGACGGT CGCGTAGTTG TTATAAAAGT GGACTTAGTC CTATAAGAAG

6651 TAATACCTGG AATGCTGTTT TCCCGGGGAT CGCAGTGGTG AGTAACCATG
ATTATGGACC TTACGACAAA AGGGCCCTTA GCGTCACCAC TCATTGGTAC

6701 CATCATCAGG AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT
GTAGTAGTCC TCATGCCCTAT TTTACGAACT ACCAGCCTTC TCCGTATTTA

6751 TCCGTCAGCC AGTTTAGTCT GACCATCTCA TCTGTAACAT CATTGGCAAC
AGGCAGTCGG TCAAATCAGA CTGGTAGAGT AGACATTGTA GTAACCGTTG

6801 GCTACCTTTG CCATGTTTCA GAAACAACTC TGGCGCATCG GGCTTCCCAT
CGATGGAAAC GGTACAAAGT CTTTGTTGAG ACCGCGTAGC CCGAAGGGTA

6851 ACAATCGATA GATTGTCGCA CCTGATTGCC CGACATTATC CCGAGCCCAT
TGTTAGCTAT CTAACAGCGT GGACTAACGG CCGTAATAG CGCTCGGGTA

XhoI

6901 TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTCGA
AATATGGGTA TATTTAGTCG TAGGTACAAC CTTAAATTAG CGCCGGAGCT

XhoI

6951 GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
CGTTCTGCAA AGGGCAACTT ATACCGAGTA TTGTGGGGAA CATAATGACA

7001 TTATGTAACC AGACAGTTT ATTGTTCAAG ATGATATATT TTTATCTTGT
AATACATTCT TCTGTCAAAA TAACAAGTAC TACTATATAA AAATAGAACA

DraIII

```
7051  GCAATGTAAC ATCAGAGATT TTGAGACACA ACGTGGCTTT CCCCCCCCCC
      CGTTACATTG TAGTCTCTAA AACTCTGTGT TGCACCGAAA GGGGGGGGGG
.....
7101  CCATTATTGA AGCATTATC AGCGTTATTG TCTCATGAGC GGATACATAT
      GGTAAATACT TCGTAAATAG TCCCAATAAC AGAGTACTCG CCTATGTATA
.....
7151  TTGAATGTAT TTAGAAAAAT AAACAATATG GGGTTCCGCG CACATTTCCT
      AACTTACATA AATCTTTTTA TTTGTTTATC CCAAGGCGC GTGTAAAGGG
.....
7201  CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC
      GCTTTTCACG GTGGACTGCA GATTCTTTGG TAATAATAGT ACTGTAATTG
.....
7251  CTATAAAAAAT AGGCGTATCA CGAGGCCCTT TCGTC
      GATATTTTTA TCCGCATAGT GCTCCGGGAA AGCAG
.....
```

pVR 1012-SGP(Z)

General Description

DNA pVR 1012-SGP(Z)
 Local object
 Created: 09/14/98 04:29PM
 Last Modified: 09/15/98 04:50PM
 length: 7272 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

DraIII: 1 site CACNNNGTG
 GTGNNNCAC

HindIII: 1 site AAGCTT
 TTCGAA

HpaI: 1 site GTTAAC
 CAATTG

KpnI: 1 site GGTACC
 CCATGG

NotI: 1 site GCGGCCGC
 CGCCGGCG

PmlI: 1 site CACGTG
 GTGCAC

PvuI: 1 site CGATCG
 GCTAGC

SacII: 1 site CCGCGG
 GGCGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATGG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATGC
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4289 End: 4841

Kan^r

Start: 6337 End: 6959 (Complementary)

Misc_feature (2 signals)

Sequence Listing ID No: 4

WO 99/32147

35

PCT/US98/27364

CMV enhancer

Start: 248 End: 885

SGP(Z)

Start: 1870 End: 4288

Annotations

```

1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
   AGCGCGCAAA GCCACTACTG CCACTTTGG AGACTGTGTA CGTCGAGGGC
.....
51  GACACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
   CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC
.....
101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
   AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC
.....

```

NdeI

```

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
   GCCGTAGTCT CGTCTAACAT GACTCTCAGC TCGTATACGC CACACTTTAT
.....
201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGCCCA
   GCGGTGTCTA CGCATTCCCT TTTTATGGCG TAGTCTAACC GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
   AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
   AGTTTGTAAT GCGCGTACAA CTGTAATAA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
   TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGCGGCAA
.....
401 ACATAACTTA CCGTAAATGG CCCGCCCTGG TGACCGCCCA ACGACCCCCG
   TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTC CATAGTAACG CCAATAGGGA
   GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTATATCCCT
.....
501 CTTCCTATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
   GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTGG ACGGGTGAAC
.....

```

NdeI

```

551 GCAGTACATC AAGTGATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
   CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT
.....
601 TGACGGTAAA TGCCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
   ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGAATACCC
.....

```

NcoI

```

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
   TCAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

```

NcoI

```

701 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
   CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG
.....
751 ACCGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
   TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAA
.....
801 GGCACCAAAA TCAACGGGAC TTCCAAAAT GTCGTAACAA CTCCGCCCCA
   CCGTGTTTTT AGTTGCCCTG AAAGGTTTA CAGCATTTG GAGCGGGGT
.....

```

851 TTGACGCAAA TGGGCGCTAG GCCTGTACGG TGGGAGGTCT ATATAAGCAG
 AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

 901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
 TCGAGCAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
 AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTGCGAGGC GCCGCCCCCT

 1001 CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
 GCCACGTAAC CTTGCCCTA AGGGGCACGG TTCTCACTGC ATTATGGCG

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
 GATATCTGAG ATATCCGTGT GGGGAACCG AGAATACGTA CGATATGACA

 1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
 AAAACCGAAC CCCGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

 1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
 ATATCGAATC GGTATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

 1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
 ATAACCACTG CTATGAAAGG TAATGATTAG CTATTGTACC GAGAAACGGT

 1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTTC AGAGACTGAC
 GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

 1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTAITA TTTACAAATT
 TGCCCTGAGAC ATAAAAATGT CCTACCCAG GGTAAATAAT AAATGTTAA

 1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTMT ATTAAACATA
 CTGTATATGT TGTTCGGCA GGCGGCACGG GCGTCAAAA TAATTTGTAT

 1401 GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
 CGCACCCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

 1451 TCTCCGGTAG CGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
 AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

 1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCTAACAG TGGAGGCCAG
 TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

 1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
 TGAATCCGTG TCGTGTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

 1601 TCGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
 ACCGCCATCC CATAACAGA CTTTACTCG CACCTCTAAC CCGAGCGTGC

 1651 GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
 CGACTGCGTC TACCTTCTGA ATTCCGTGCG CGTCTTCTTC TACGTCCGTC

 1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTCC
 GACTCAACAA CATAAGACTA TTCTCACTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTAAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCCGG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
CGCGGGTGGT CTGTATTATC GACTGTCTGA TTCTCTGACA AGGAAAGGTA

NcoIPmlIEcoRVNotI

1851 GGGCTTTTTC TGCAGTCACC GTCGTGACA CGTCTGATCA CATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

NotI XbaI

1901 CCGCTCTAGA CCAGCGCCTT GGATCGAATT GATGAAGATT AAGCCGACAG
GGCTAGATCT GGTCCGCGGA CCTAGCTTAA CTACTTCTAA TTCGGCTGTC

1951 TGAGCGTAAT CTTTCATCTCT CTTAGATTAT TTGTTTCCA GAGTAGGGGT
ACTCGCATTA GAAGTAGAGA GAATCTAATA AACAAAAGGT CTCATCCCCA

2001 CSTCAGGTCC TTTTCAATCG TGTAACCAA ATAACTCCA CTAGAAGGAT
CCAGTCCAGG AAAAGTTAGC ACATTGGTTT TATTTGAGGT GATCTTCTTA

2051 ATTGTGGGCG AACACACAA TGGGCGTTAC AGGAATATTG CAGTTACCTC
TAACACCCCG TTGTGTGTT ACCCGCAATG TCCTTATAAC GTCAATGGAG

2101 GTGATCGATT CAAGAGGACA TCATTCTTTC TTGCGTAAAT TATCCTTTTC
CACTAGCTAA GTTCTCCTGT AGTAACAAAG AAACCCATTA ATAGGAAAAG

2151 CAAAGAACAT TTTCCATCCC ACTTGGAGTC ATCCACAATA GCACATTACA
GTTTCTTGTA AAAGGTAGGG TGAACCTCAG TAGGTGTTAT CGTGTAAATG

2201 GGTTAGTGAT GTCGACAAAC TAGTTTGTGCG TGACAAACTG TCATCCACAA
CCAATCACTA CAGCTGTTTG ATCAACAGC ACTGTTTGAC AGTAGGTGTT

2251 ACCAATTGAG ATCAGTTGGA CTGAATCTCG AAGGGAATGG AGTGGCAACT
TAGTTAATC TAGTCAACCT GACTTAGAGC TTCCCTTACC TCACCGTTGA

2301 GACGTGCCAT CTGCAACTAA AAGATGGGGC TTCAGGTCCG GTGTCCCACC
CTGCACGGTA GACGTTGATT TTCTACCCCG AAGTCCAGGC CACAGGGTGG

2351 AAGGTTGGTC AATTATGAAG CTGGTGAATG GGCTGAAAAC TCCTACAATC
TTTCCACCAG TTAATACTTC GACCACTTAC CCGACTTTTG ACGATGTTAG

2401 TTGAAATCAA AAAACCTGAC GGGAGTGAGT GTCTACCAGC AGCGCCAGAC
AACTTTAGTT TTTTGGACTG CCTTCACTCA CAGATGGTGG TCGCGGTCTG

2451 CGGATTCGGG GCTTCCCCCG GTGCCGGTAT GTGCACAAAG TATCAGGAAC
CCCTAAGCCC CCAAGGGGGC CACGGCCATA CACGTGTTTC ATAGTCCTTG

2501 GGCACCGTGT GCCGGAGACT TTCCCTTCCA TAAAGAGGGT GCTTCTTCC
CCCTGGCACA CGCCCTCTGA AACGGAAGGT ATTTCTCCCA CGAAAGAAGG

2551 TGTATGATCG ACTTGCTTCC ACAGTTATCT ACCGAGGAAC GACTTTCGCT
ACATACTAGC TGAACGAAGG TGTCAATAGA TGGCTCCTTG CTGAAAGCGA

2601 GAAGGTCTCG TTGCATTTCT GATACTGCCC CAAGCTAAGA AGGACTTCTT
CTTCCACAGC AACGTAAAGA CTATGACGGG GTTCGATTCT TCCTGAAGAA

2651 CAGCTCACAC CCCTTGAGAG AGCCGGTCAA TGCAACGGAG GACCCGTCTA
GTCCAGTGTG GGGAACTCTC TCGGCCAGTT ACCTTGCCCTC CTGGGCAGAT

EcoRV

2701 GTGGCTACTA TTCTACCACA ATTAGATATC AGGCTACCGG TTTTGGAACC
CACCGATGAT AAGATGGTGT TAATCTATAG TCCGATGGCC AAAACCTTGG

2751 AATGAGACAG AGTACTTGTG CGAGGTTGAC AATTTGACCT ACGTCCAACCT
TTACTCTGTC TCATGAACAA GCTCCAACCTG TTAAACTGGA TGCAGGTTGA

2801 TGAATCAAGA TTCACACCAC AGTTTCTGCT CCAGCTGAAT GAGACAATAT
ACTTAGTTCT AACTGTGCTG TCAAAGACGA GGTCGACTTA CTCTGTTATA

2851 ATACAAGTGG GAAAAGGAGC AATACCACGG GAAACTAAT TTGGAAGGTC
TATGTTTACC CTTTCTCTCG TTATGGTGCC CTTTGATTA AACCTTCCAG

2901 AACCCCGAAA TTGATACAAC AATCGGGGAG TGGCCCTTCT GGGAACTAA
TTGGGGCTTT AACTATGTG TTAGCCCTC ACCCGGAAGA CCCTTTGATT

2951 AAAAACCTCA CTAGAAAAAT TCGCAGTGAA GAGTTGTCTT TCACAGTTGT
TTTTTGGAGT GATCTTTTAA AGCGTCACTT CTCACAGAA AGTGTCAACA

3001 ATCAAACGGA GCCAAAAACA TCAGTGGTCA GAGTCCGGCG CGAACTTCTT
TAGTTTCCCT CGCTTTTGT AGTCACCAGT CTCAGGCCGC GCTTGAAGAA

3051 CCGACCCAGG GACCAACACA ACAACTGAAG ACCACAAAAT CATGCCCTCA
GGCTGGGTCC CTGTTGTGT TGTGACTTC TGGTGTTTA GTACCGAAGT

3101 GAAAAATTCCT CTGCAATGGT TCAAGTGCAC AGTCAAGGAA GGAAGCTGC
CTTTTAAGGA GACGTTACCA AGTTCACGTG TCAGTTCCTT CCTTCGACG

3151 AGTGTCCAT CTAACAACCC TTGCCACAAT CTCCACGAGT CCCCATCCC
TCACAGCGTA GATTGTTGGG AACGGTGTAA GAGGTGCTCA GGGGTTAGGG

3201 TCACAACCAA ACCAGGTCCG GACAACAGCA CCCATAATAC ACCCGTGTAT
AGTGTGTTGTT TGGTCCAGGC CTGTTGTCTG GGGTATTATG TGGGCACATA

3251 AAACCTTGACA TCTCTGAGGC AACTCAAGTT GAACAACATC ACCGCAGAAC
TTTGAACGT AGAGACTCCG TTGAGTTCAA CTTGTTGTAG TGGCGTCTTG

3301 AGACAACGAC AGCACAGCCT CCGACACTCC CTCTGCCACG ACCGCAGCCG
TCTGTTGCTG TCGTGTCCGA GGCTGTGAGG GAGACGGTGC TGGCGTCGGC

3351 GACCCCAAAA AGCAGAGAAC ACCAACACGA CCAAGAGCAC TGACTTCCTG
CTGGGGGTTT TCGTCTCTTG TGGTTGTGCT CGTTCTCGTG ACTGAAGGAC

3401 GACCCCGCCA CCACAACAAG TCCCCAAAAC CACAGCGAGA CCGCTGGCAA
CTGGGGCGGT GGTGTTCTTC AGCGTTTTG GTGTCGCTCT GCGACCGTT

3451 CAACAACACT CATACCAAG ATACCGGAGA AGAGAGTGCC AGCAGCGGA
GTTGTTGTGA GTAGTGGTTC TATGGCCTCT TCTCTACCG TCGTCGCCCT

```

3501 AGCTAGGCTT AATTACCAAT ACTATTGCTG GAGTCGCAGG ACTGATCACA
    TCGATCCGAA TTAATGGTTA TGATAACGAC CTCAGCGTCC TGACTAGTGT
.....
3551 GCGGGGAGAA GAACTCGAAG AGAAGCAATT GTCAATGCTC AACCCTAATG
    CCGCCCTCTT CTTGAGCTTC TCTTCGTAA CAGTTACGAG TTGGGTTTAC
.....
3601 CAACCCTAAT TTACATTACT GGACTACTCA GGATGAAGGT GCTGCAATCG
    CTTGGGATTA AATGTAATGA CCTGATGAGT CCTACTTCCA CGACGTTAGC
.....
3651 GACTGGCCTG CATAACATAT TTCGGGCCAG CAGCCGAGGG AATTACATA
    CTGACCGGAC CTATGGTATA AAGCCCCGTC GTCCGGCTCC TTAATGTAT
.....
3701 GAGGGGCTAA TGCACAATCA AGATGGTTTA ATCTGTGGGT TGAGACAGCT
    CTCCCCGATT ACGTGTACT TCTACCAAAT TAGACACCCA ACTCTGTGCA
.....
3751 GGCCAACGAG ACGACTCAAG CTCTCAACT GTTCCTGAGA GCCACAACCTG
    CCGGTTGCTC TGCTGAGTTC GAGAAGTTGA CAAGGACTCT CGGTGTTGAC
.....
3801 ACCTACGCAC CTTTTCAATC CTCAACCGTA AGGCAATTGA TTTCTTGCTG
    TCGATCCGTG GAAAGTTAG GAGTTGGCAT TCCGTAACT AAAGAACGAC
.....
3851 CAGCGATGGG GCGGCACATG CCACATTCTG GGACCGGACT GCTGTATCGA
    GTCGCTACCC CGCCGTGTAC GGTGTAAGAC CCTGGCCTGA CGACATAGCT
.....
3901 ACCACATGAT TGGACCAAGA ACATAACAGA CAAAATTGAT CAGATTATTC
    TGGTGTACTA ACCTGGTTCT TGTATTGTCT GTTTAACTA GTCTAATAAG
.....
3951 ATGATTTTGT TGATAAAACC CTTCCGGACC AGGGGGACAA TGACAATTGG
    TACTAAAACA ACTATTTTGG GAAGGCCTGG TCCCCCTGTT ACTGTAAACC
.....
4001 TGGACAGGAT CGAGACAATG GATACCGGCA GGTATTGGAG TTACAGGCCT
    ACCTGTCTTA CCTCTGTTAC CTATGGCCGT CCATAACCTC AATGTCCGCA
.....
4051 TATAATTGCA GTTATCGCTT TATTCTGTAT ATGCAAAATT GTCTTTTAGT
    ATATTAACGT CAATAGCGAA ATAAGACATA TACGTTTAA CAGAAAATCA
.....
4101 TTTTCTTCAG ATTGCTTCAT GGAAAAGCTC AGCCTCAAAT CAATGAAACC
    AAAAGAAGTC TAACGAAGTA CCTTTTCGAG TCGGAGTTTA GTTACTTTGG
.....
4151 AGGATTTAAT TATATGGATT ACTTGAATCT AAGATTACTT GACAAATGAT
    TCCTAAATTA ATATACCTAA TGAACCTAGA TTCTAATGAA CTGTTTACTA
.....
4201 AATATAATAC ACTGGAGCTT TAAACATAGC CAATGTGATT CTAACCTCCT
    TTATATTATG TGACCTCGAA ATTGTATCG GTTACACTAA GATTGAGGAA
.....
4251 TAAACTCACA GTTAATCATA AACAAGGTTT CGAATTGATC TGCTGTGCCT
    ATTTGAGTGT CAATTAGTAT TTGTTCCAAA CCTTAACCTAG ACGACACGGA
.....
4301 TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCCGTGC CTTCCTTGAC
    AGATCAACGG TCGGTAGACA ACAAACGGGG AGGGGGCACC GAAGGAACCTG
.....
4351 CTTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
    GGACCTTCCA CGGTGAGGGT GACAGGAAAG GATTATTTTA CTCCTTTAAC
.....
4401 CATCGCATTG TCTGAGTAGG TGTCAATCTA TTCTGGGGGG TGGGGTGGGG
    GTACCGTAAC AGACTCATCC ACAGTAAGAT AAGACCCCCC ACCCCACCCC
.....

```


SphI

4451 CAGCACAGCA AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
GTCGTGTCGT TCCCCCTCCT AACCCCTTCTG TTATCGTCCG TACGACCCCT

KpnI

4501 TCGGGTGGGC TCTATGGGTA CCCAGGTGCT GAAGAATTGA CCCGGTTCCT
ACGCCACCCG AGATACCCAT GGTCCACGA CTTCTTAAC TGGCCAAGGA

4551 CCTGGGCCAG AAAGAAGCAG GCACATCCCC TTCTCTGTGA CACACCCTGT
CGACCCGGTC TTTCTTCGTC CGTCTAGGGG AAGAGACACT GTGTGGGACA

4601 CCACGCCCTT GGTTCCTTAGT TCCAGCCCCA CTCATAGGAC ACTCATAGCT
GGTGGCGGCA CCAAGAATCA AGGTGGGGGT GAGTATCTG TGAGTATCGA

4651 CAGGAGGGCT CCGCCTTCAA TCCACCCGCG TAAAGTACTT GGAGCGGTCT
GTCCTCCCGA GCGGAAGTT AGGCTGGGCG ATTTTCATGAA CCTCGCCAGA

4701 CTCCCTCCCT CATCAGCCCA CCAAAACAAA CCTAGCCTCC AAGAGTGGGA
GAGGGAGGGA GTAGTCGGGT GGTTCGGTTT GGATCGGAGG TTCACACCT

4751 AGAAATTAAA CCAAGATAGG CTATTAAGTC CAGAGGGAGA GAAATGCCT
TCATTAAATT CGTCTATCC GATAATTCAC GTCTCCCTCT CTTTACGGA

4801 CCAACATGTG AGGAAGTAAT GAGAGAAATC ATAGAATTTC TTCCGCTCC
GGTTGTACAC TCCTTCATTA CTCTCTTAG TATCTTAAAG AAGGCGAAGG

4851 TCGCTCACTG ACTCGCTGCG CTCGCTCGTT CGGCTGCGGC GAGCGGTATC
AGCGAGTGC TGAGCGACGC GAGCCAGCAA GCCGACGCCG CTCGCCATAG

4901 AGCTCACTCA AAGCGGGTAA TACGGTTATC CACAGAATCA GGGGATAACG
TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT CCCCTATTGC

4951 CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAA
GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATT

5001 AAGGCCGCGT TGCTGGCGTT TTCCATAGG CTCCGCCCCC CTGACGAGCA
TTCCGGCGCA ACGACCCCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT

5051 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
AGTGTTTTFA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC TGTCTGATA

5101 AAAGATACCA GCGCTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT
TTTCTATGGT CCCCRAAGGG GGACCTTCCA GGGAGCACGC GAGAGGACAA

5151 CCGACCCCTG CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCCGGGAG
GGCTGGGACG CGGAATGGCC TATGGACAGG CGGAAAGAGG GAAGCCCTTC

5201 CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA AGCCACATCC

5251 TCGTTGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCGGGCTG

5301 CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCACC CGGTAAGACA
GGGACCGGGA ATAGGCCATT GATACCAGAA CTCAGGTTGG GCCATTCTGT

```

5351  CGACTTATCG  CCACTGCCAG  CAGCCACTGG  TAACAGGATT  AGCAGAGCGA
      GCTGAATAGC  GGTGACCGTC  GTCGGTGACC  ATTGTCTCTA  TCGTCTCGCT
.....
5401  GGTATGTAGG  CGGTGCTACA  GAGTTCTTGA  AGTGGTGGCC  TAACTACGGC
      CCATACATCC  CCCACGATGT  CTCAAGAACT  TCACCACCGG  ATTGATGCCG
.....
5451  TACACTAGAA  GGACAGTATT  TGGTATCTGC  GCTCTGCTGA  AGCCAGTTAC
      ATGTGATCTT  CCTGTCATAA  ACCATAGACC  CGAGACCACT  TCGGTCAATG
.....
5501  CTTTCGGAAA  AGAGTTGGTA  CCTCTTGATC  CGGCAAACAA  ACCACCGCTG
      GAAGCCTTTT  TCTCAACCAT  CGAGAACTAG  GCCGTTTGTT  TGGTGCCGAC
.....
5551  GTAGCGGTGG  TTTTMTTGT  TGCAAGCAGC  AGATTACCGG  CAGAAAAAAA
      CATCGCCACC  AAAAAACAA  ACGTTCGTCC  TCTAATGCCG  GTCTTTTMTT
.....
5601  GGATCTCAAG  AAGATCCTTT  GATCTTTTCT  ACGGGGTCTG  ACGCTCAGTG
      CCTAGAGTTC  TTCTAGGAAA  CTAGAAAAGA  TGCCCCAGAC  TGCGAGTCAC
.....
5651  GAACGAAAAC  TCACGTTAAG  GGATTTTGGT  CATGAGATTA  TCAAAAAGGA
      CTTGCTTTTG  AGTGCAATTC  CCTAAAACCA  GTACTCTAAT  AGTTTTCTCT
.....
5701  TCTTCACCTA  GATCCTTTTA  AATTAAAAAT  GAAGTTTTAA  ATCAATCTAA
      AGAAGTGGAT  CTAGGAAAAT  TTAATTTTAA  CTTCAAAATT  TAGTTAGATT
.....
5751  AGTATATATG  AGTAAACTTG  GTCTGACAGT  TACCAATGCT  TAATCAGTGA
      TCATATATAC  TCATTTGAAC  CAGACTGTCA  ATGGTTACGA  ATTAGTCACT
.....
5801  GGCACCTATC  TCAGCGATCT  GTCTATTTCT  TTCATCCATA  GTTGCCTGAC
      CCGTGGATAG  AGTCGCTAGA  CAGATAAAGC  AAGTAGGTAT  CAACGGACTG
.....
5851  TCCGGGGGGG  GGGGGCGCTG  AGGTCTGCCT  CGTGAAGAAG  GTGTTGCTGA
      AGGCCCCCCC  CCCCCCGGAC  TCCAGACGGA  GCACTTCTTC  CACAACGACT
.....
5901  CTCATACCAG  GCCTGAATCG  CCCCATCATC  CAGCCAGAAA  GTGAGGGAGC
      GAGTATGGTC  CGGACTTAGC  GGGGTAGTAG  GTCGGTCTTT  CACTCCCTCG
.....
5951  CACGGTTGAT  GAGAGCTTTG  TTGTAGGTGG  ACCAGTTGGT  GATTTTGAAC
      GTGCCAACTA  CTCTCGAAAC  AACATCCACC  TGGTCAACCA  CTAAAACCTG
.....
6001  TTTTGCTTTG  CCACGGAACG  GTCTGCGTTG  TCGGGAAGAT  GCGTGATCTG
      AAAACGAAAC  GGTGCCTTGC  CAGACGCAAC  AGCCCTTCTA  CGCACTAGAC
.....
6051  ATCCTTCAAC  TCAGCAAAAG  TTGGATTTAT  TCAACAAAGC  CGCCGTCCCG
      TAGGAAGTTG  AGTCGTTTTT  AAGCTAAATA  AGTTGTTTCG  GCGGCAGGGC
.....
6101  TCAAGTCAGC  GTAATGCTCT  GCCAGTGTTA  CAACCAATTA  ACCAATTCTG
      AGTTCAGTCG  CATTACGAGA  CGGTCACAAT  GTTGGTTAAT  TGOTTAAGAC
.....
6151  ATTAGAAAAA  CTCATCGAGC  ATCAAATGAA  ACTGCAATTT  ATTCATATCA
      TAATCTTTTT  GAGTAGCTCG  TAGTTTACTT  TGACCTTAAA  TAAGTATAGT
.....
6201  GGATTATCAA  TACCATATTT  TTGAAAAAGC  CGTTTCTGTA  ATGAAGCAGA
      CCTAATAGTT  ATGGTATAAA  AACTTTTTCG  GCAAAGACAT  TACTTCCTCT
.....
6251  AAAGTCACCG  AGGCAAGTCC  ATAGGATGGC  AAGATCCTGG  TATCGGTCTG
      TTTGAGTGGC  TCCGTCAAGG  TATCCTACCG  TTCTAGGACC  ATAGCCAGAC
.....

```

6301 CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
GCTAAGGCTG AGCAGGTTGT AGTTATGTTG GATAATTAAA GGGGAGCAGT

6351 AAAATAAGGT TATCAAGTGA GAAATCACCA TGAGTGACGA CTGAATCCGG
TTTTATTCCA ATAGTTCACT CTTAGTGGT ACTCACTGCT GACTTAGGCC

HindIII

6401 TGAGAATGGC AAAAGCTTAT GCATTTCTTT CCAGACTTGT TCAACAGGCC
ACTCTTACCG TTTTCGAATA CGTAAAGAAA GGTCTGAACA AGTTGTCCGG

6451 AGCCATTACG CTCGTCATCA AAATCACTCG CATCAACCAA ACCGTTATTC
TCGGTAATGC GAGCAGTAGT TTTAGTGAGC GTAGTTGGTT TGGCAATAAG

PvuI

6501 ATTCTGTGATT GCGCCTGAGC GAGACGAAAT ACGCGATCGC TGTTAAAGG
TAAGCACTAA CGCGGACTCG CTCGTCTTTA TCGGCTAGCG ACAATTTTCC

6551 ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC ACTGCCAGCG
TGTTAATGTT TGTCCTTAGC TTACGTTGGC CGCGTCTTTG TGACGGTCCG

6601 CATCAACAAT ATTTTCACCT CAATCAGGAT ATTCTTCTAA TACCTGGAAT
GTAGTTGTTA TAAAAGTGGA CTTAGTCCTA TAAGAAGATT ATGGACCTTA

6651 GCTGTTTCC CGGGGATCGC AGTGGTGAGT AACCATGCAT CATCAGGAGT
CGACAAAAGC GCCCTTAGCG TCACCACTCA TTGGTACGTA GTAGTCCTCA

6701 ACGGATAAAA TGCTTGATGG TCGGAAGAGG CATAAATTCC GTCAGCCAGT
TCCCTATTTT ACGAACTACC AGCCTTCTCC GTATTTAAGG CAGTCGGTCA

6751 TTAGTCTGAC CATCTCATCT GTAACATCAT TGGCAACGCT ACCTTTGCCA
AATCAGACTG GTAGAGTAGA CATTGTAGTA ACCGTTGCGA TGGAAACGGT

6801 TGTTTCAGAA ACAACTCTGG CGCATCGGGC TTCCCATACA ATCGATAGAT
ACAAAGTCTT TGTTGAGACC GCGTAGCCCG AAGGGTATGT TAGCTATCTA

6851 TGTCGCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA TACCCATATA
ACAGCGTGGA CTAACGGGCT GTAATAGCGC TCGGGTAAT ATGGGTATAT

XhoI

6901 AATCAGCATC CATGTTGGAA TTTAATCGCG GCCTCGAGCA AGACGTTTCC
TTAGTCGTAG GTACAACCTT AAATTAGCGC CGGAGCTCGT TCTGCAAAGG

6951 CGTTGAATAT GGCTCATAAC ACCCCTTGTA TTACTGTTTA TGTAAGCAGA
GCAACTTATA CCGAGTATTG TGGGGAACAT AATCACAAT ACATTCTGCT

7001 CAGTTTTATT GTTCATGATG ATATATTTTT ATCTGTGCGA ATGTAACATC
GTCAAAATAA CAAGTACTAC TATATAAAAA TAGAACACGT TACATTGTAG

DraIII

7051 AGAGATTTTG AGACACAACG TGCCTTTCCC CCCCCCCCCA TTATTGAAGC
TCTCTAAAAC TCTGTGTTGC ACCGAAAGGG GGGGGGGGGT AATAACTTCG

7101 ATTTATCAGG GTTATTGTCT CATGAGCGCA TACATATTTG AATGTATTTA
TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

7151 CAAAAATAAA CAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTCCCAC
CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG

7201 CTGACGTCTA AGAAACCATT ATTATCATGA CATTACCTA TAAAAATAGG
GACTGCAGAT TCTTTGGTAA TAATAGTACT GTAATTGGAT ATTTTATCC

7251 CGTATCACGA GGCCCTTTTCG TC
GCATAGTCCT CCGGGAAAGC AG